

CORRÉLATIONS

et tout ti kwan ti.

Rappels sur les fonctions de corrélation

Les fonctions de corrélation permettent de mesurer une certaine « ressemblance » statistique entre deux signaux aléatoires. On définit la fonction d'autocorrélation $R_{xx}(\tau)$, dans le cas d'un processus stationnaire, par

$$R_{xx}(\tau) = E[X(t) X^*(t-\tau)],$$

et la fonction d'intercorrélation par

$$R_{xy}(\tau) = E[X(t) Y^*(t-\tau)].$$

Rappelons quelles sont les propriétés élémentaires des fonctions de corrélation :

- R_{xx} est une fonction paire,
- $R_{xx}(0) = E[|X(t)|^2] =$ Puissance de x ,
- R_{xx} est maximale en zéro,
- $R_{xx}(\tau)$ est une fonction définie non négative, ce qui signifie en particulier que sa transformée de Fourier est ≥ 0 . (Théorème de Bochner).
- $R_{xy}(\tau)$ est à symétrie hermitienne :
$$R_{xy}(-\tau) = R_{yx}^*(\tau)$$

Lorsque les processus sont ergodiques (et stationnaire), on peut substituer une moyenne temporelle aux moyennes d'ensemble, et, par exemple :

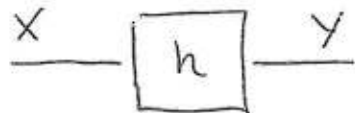
$$R_{xy}(\tau) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t) Y^*(t-\tau) dt$$

$$R_{xy}(k) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{-\frac{N}{2}}^{\frac{N}{2}} X(n) Y^*(k-n)$$

pour des signaux à temps discret.

Corrélation et filtrage

Pour un signal aléatoire filtré selon



on dispose de la fameuse formule des interférences.

$$R_{YY} = h * h^{(-)} * R_{XX}$$

$$\text{ou } h^{(-)}(u) = h^*(-u)$$

$$\text{et } R_{YX} = h * R_{XX}.$$

Le théorème de Wiener-Kintchine-Einstein indique que la densité spectrale de puissance est la TF de la fonction d'autocorrélation.

$$S_{XX} \rightleftharpoons R_{XX}$$

A partir de la formule des interférences, on en déduit que

$$S_{YY}(f) = |H(f)|^2 S_{XX}(f)$$

en utilisant le fait que $TF\{h^{(-)}\} = H^*$.

Dans le cas d'un bruit blanc X , la densité spectrale est constante, disons de valeur σ^2 , et

$$S_{YY}(f) = \sigma^2 |H(f)|^2.$$

L'un des intérêts des fonctions de corrélation est qu'elles permettent la description complète d'un processus gaussien. En effet, si $\underline{x}(n)$ est un vecteur gaussien, sa densité de probabilité est donnée par

$$p_{\underline{x}}(\underline{x}) = \frac{1}{(2\pi |\underline{R}_x|^2)^{\frac{N}{2}}} \exp_{\frac{1}{2}}(\underline{x} - \underline{m})^{\dagger} \underline{R}_x^{-1} (\underline{x} - \underline{m})$$

où $\underline{m} = E[\underline{x}]$

et $\underline{R}_x = E[(\underline{x} - \underline{m})(\underline{x} - \underline{m})^{\dagger}]$ est la matrice de corrélation (ou covariance)

lorsque le vecteur aléatoire est construit comme la suite de N échantillons

$$\underline{x}(n) = [X(n) \ X(n-1) \ \dots \ X(n-N+1)]^T$$

La matrice de corrélation s'écrit (dans le cas centré $\underline{m} = \underline{0}$ pour alléger l'écriture)

$$E \left\{ \begin{array}{c} \left[\begin{array}{c} X(n) \\ X(n-1) \\ \vdots \\ X(n-N+1) \end{array} \right] \left[X(n)^* \ X(n-1)^* \ \dots \ X(n-N+1)^* \right] \end{array} \right\}$$

$$\underline{R}_{\underline{x}x} = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(N-1) \\ R_{xx}(-1) & R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(N-2) \\ \vdots & & & & \vdots \\ R_{xx}(-N+1) & \dots & R_{xx}(1) & R_{xx}(0) \end{bmatrix}$$

Qui est une matrice de Toeplitz (on retrouve les mêmes termes sur chacune des diagonales) et à symétrie hermitienne (car $R_{xx}(k) = R_{xx}(k)^*$) :

$$\underline{R}_{xx} = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(N-1) \\ R_{xx}(1)^* & R_{xx}(0) & & \\ \vdots & & \ddots & \\ R_{xx}(N-1)^* & \dots & & R_{xx}(0) \end{bmatrix}$$

avec $\underline{R}_{xx} = \underline{R}_{xx}^+$

Ainsi, la densité de probabilité à N instants d'un processus gaussien est entièrement fixée par la donnée de la moyenne et de la fonction d'autocorrélation.

Évidemment tel n'est pas le cas des densités de probas en général, mais la présence de nombreux systèmes gaussiens, la simplicité et l'efficacité des calculs à l'ordre 2 (notion d'espace de Hilbert, etc) et les interprétations énergétiques conduisent souvent à se limiter à l'ordre 2.

Dérivation par rapport à un complexe, par rapport à un vecteur.

Une fonction y d'une variable complexe ne peut être dérivée par rapport à un complexe z que si elle est analytique et vérifie la condition de Cauchy-Riemann

$$\frac{dy}{dz_I} = j \frac{dy}{dz_R}$$

En effet, une fonction $y(z) / y(z+d) = y(z) + dy'(z) + O(d^2)$ n'existe pas nécessairement pour tout $d \in \mathbb{C}$.

En particulier, la conjugaison complexe n'est pas une opération analytique, et les fonctions mettant en jeu une conjugaison ne peuvent être dérivées.

On peut définir une dérivée généralisée:

$$\frac{dy}{dz} = \left(\frac{dy}{dz_R} - j \frac{dy}{dz_I} \right) / 2$$

et la dérivée conjuguée est

$$\frac{dy}{dz^*} = \left(\frac{dy}{dz_R} + j \frac{dy}{dz_I} \right) / 2$$

Si y est analytique, on aura

$$\frac{dy}{dz} = \frac{dy}{dz_R} = -j \frac{dy}{dz_I}$$

$$\text{et } \frac{dy}{dz^*} = 0 \text{ et } \frac{dy^*}{dz} = 0$$

De manière générale, on a également

$$\frac{dy^*}{dz^*} = \left(\frac{dy}{dz} \right)^*$$

NB - Bien sûr, on a $\frac{dz}{dz} = 1$ - D'autre part $\frac{dz}{dz^*} = 0$

On définit la dérivée par rapport à un vecteur comme l'application de l'opérateur

$$\frac{d}{d\underline{z}} = \begin{bmatrix} \frac{d}{d\underline{z}(1)} \\ \vdots \\ \frac{d}{d\underline{z}(p)} \end{bmatrix}$$

avec $\underline{z}(k)$ la k^e composante de \underline{z} .

La dérivée d'une fonction scalaire est alors un vecteur, et la dérivée d'un vecteur ligne est une matrice

$$\frac{df}{d\underline{z}} = \begin{bmatrix} \frac{df(\underline{z})}{d\underline{z}(1)} \\ \vdots \\ \frac{df(\underline{z})}{d\underline{z}(p)} \end{bmatrix}$$

$$\frac{d\underline{y}^T}{d\underline{z}} = \begin{bmatrix} \frac{dy(1)}{d\underline{z}(1)} & \dots & \frac{dy(q)}{d\underline{z}(1)} \\ \vdots & & \vdots \\ \frac{dy(1)}{d\underline{z}(p)} & \dots & \frac{dy(q)}{d\underline{z}(p)} \end{bmatrix}$$

Dans le cas complexe, on définit le gradient $\nabla_{\underline{z}}$ d'une fonction à valeurs réelles par

$$\nabla_{\underline{z}} [f(\underline{z})] = 2 \frac{df(\underline{z})}{d\underline{z}^*} = 2 \left(\frac{df(\underline{z})}{d\underline{z}} \right)^*$$

On montre que

$$\nabla_{\underline{z}} [f(\underline{z})] = \underline{0} \text{ pour les extrema de } f(\underline{z})$$

et $\nabla_{\underline{z}}$ est \perp aux surfaces $f(\underline{z}) = \text{constante}$.

Règle du produit.

Si y_3 est un scalaire (complexe)

$$\frac{d(y_3)}{dz} = y \frac{dz}{dz} + \frac{y}{z} \frac{dy}{dz} \text{ reste vraie.}$$

Dérivées de fonctions analytiques (ou réelles)

$$\frac{d}{d\underline{x}} (\underline{x}^T \underline{a}) = \frac{d}{d\underline{x}} (\underline{a}^T \underline{x}) = \underline{a}$$

$$\frac{d}{d\underline{x}} (\underline{x}^T \underline{A}) = \underline{A}$$

$$\frac{d}{d\underline{x}} (\underline{A}\underline{x} + \underline{b})^T \underline{C} (\underline{D}\underline{x} + \underline{e}) = \underline{A}^T \underline{C} (\underline{D}\underline{x} + \underline{e}) + \underline{D}^T \underline{C}^T (\underline{A}\underline{x} + \underline{b})$$

$$\begin{aligned} d \underline{x}^T \underline{C} \underline{x} &= (\underline{C} + \underline{C}^T) \underline{x} \\ &= 2 \underline{C} \underline{x} \quad \text{si } \underline{C} = \underline{C}^T \end{aligned}$$

$$\begin{aligned} \frac{d}{d\underline{x}} (\underline{A}\underline{x} + \underline{b})^T \underline{C} (\underline{A}\underline{x} + \underline{b}) &= \underline{A}^T (\underline{C} + \underline{C}^T) (\underline{A}\underline{x} + \underline{b}) \\ &= 2 \underline{A}^T \underline{C} (\underline{A}\underline{x} + \underline{b}) \quad \text{si } \underline{C} = \underline{C}^T \end{aligned}$$

Dérivées de fonctions non analytiques:

$$\frac{d}{d\underline{x}} (\underline{A}\underline{x} + \underline{b})^+ \underline{C} (\underline{D}\underline{x} + \underline{e}) = \underline{D}^T \underline{C}^T (\underline{A}\underline{x} + \underline{b})^*$$

$$\frac{d}{d\underline{x}} (\underline{x}^+ \underline{C} \underline{x}) = \underline{C}^T \underline{x}^*$$

$$\frac{d}{d\underline{x}} (\underline{x}^+ \underline{A}) = \underline{0}$$

$$\frac{d}{d\underline{x}} (\underline{x}^+ \underline{x}) = \underline{x}^*$$

$$\frac{d}{d\underline{x}^*} (\underline{x}^+ \underline{A}) = \underline{A}$$

Gradient:

$$\underline{\nabla}_{\underline{x}} (\underline{A}\underline{x} + \underline{b})^+ \underline{C} (\underline{D}\underline{x} + \underline{e}) = 2 \underline{D}^+ \underline{C}^+ (\underline{A}\underline{x} + \underline{b})$$

$$\underline{\nabla}_{\underline{x}} (\underline{x}^+ \underline{C} \underline{x}) = 2 \underline{C}^+ \underline{x}$$

$$\underline{\nabla}_{\underline{x}} (\underline{x}^+ \underline{x}) = 2 \underline{x}$$

Prédiction linéaire en moyenne quadratique.

Le problème de la prédiction linéaire en moyenne quadratique est de prédire une valeur future d'un signal à partir de p échantillons de ce signal - La prédiction à un pas consiste à prédire la valeur de $x(n)$ à partir de $\{x(n-1) \dots x(n-p)\}$ -

On pose

$$\hat{x}(n) = \sum_{i=1}^p \alpha_i x(n-i)$$

et il s'agit d'ajuster les coefficients $\{\alpha_i\}_{i=1:p}$ afin de rendre minimale la puissance de l'erreur de prédiction $e(n) = x(n) - \hat{x}(n)$

$$J = E[|x(n) - \hat{x}(n)|^2] = E\left[\left|x(n) - \sum_{i=1}^p \alpha_i x(n-i)\right|^2\right]$$

En posant

$$\underline{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_p]^T$$

$$\underline{x}(n) = [x(n-1), \dots, x(n-p)]^T,$$

on a $\hat{x}(n) = \underline{\alpha}^T \underline{x}(n)$ et

$$J = E[|x(n) - \underline{\alpha}^T \underline{x}(n)|^2]$$

$$= E\left[(x(n) - \underline{\alpha}^T \underline{x}(n))^* (x(n) - \underline{\alpha}^T \underline{x}(n))\right]$$

Le gradient complexe s'écrit alors, en notant que $\underline{\alpha}^T \underline{x}(n) = \underline{x}(n)^T \underline{\alpha}$,

et à partir des formules de dérivation

$$\text{avec } A = D = -\underline{x}(n)$$

$$b = e = \underline{x}(n)$$

$$\underline{\nabla}_{\underline{\alpha}} \left[\begin{matrix} \varphi \\ \xi \end{matrix} \right] = -2E \left[\underline{x}(n)^* \left(\underline{x}(n) - \underline{\alpha}^T \underline{x}(n) \right) \right]$$

ou, à partir de la règle du produit:

$$\underline{\nabla}_{\underline{\alpha}} \left[E |e(n)|^2 \right] = 2 \left(\frac{d E |e(n)|^2}{d \underline{\alpha}} \right)^* = 2E \left[\begin{matrix} e(n)^* \frac{de(n)}{d \underline{\alpha}} + \\ e(n) \frac{de(n)^*}{d \underline{\alpha}} \end{matrix} \right]$$

$e(n) = \underline{x}(n) - \underline{\alpha}^T \underline{x}(n)$ est une fonction linéaire analytique et l'on a alors

$$\frac{de(n)^*}{d \underline{\alpha}} = 0 \quad \frac{de(n)}{d \underline{\alpha}} = -\underline{x}(n).$$

Il reste donc finalement

$$\underline{\nabla}_{\underline{\alpha}} \left[\begin{matrix} \varphi \\ \xi \end{matrix} \right] = -2E \left[\underline{x}(n)^* e(n) \right].$$

À l'optimum, on a donc $\underline{\nabla}_{\underline{\alpha}} \left[\begin{matrix} \varphi \\ \xi \end{matrix} \right] = \underline{0}$

$$\underline{\nabla}_{\underline{\alpha}} \left[\begin{matrix} \varphi \\ \xi \end{matrix} \right] = -2E \left[\underline{x}(n)^* e(n) \right] = -2E \left[\underline{x}(n)^* \left(\underline{x}(n) - \underline{\alpha}^T \underline{x}(n) \right) \right]$$

Et l'on note que l'erreur de prédiction $e(n)$ est orthogonale à $\underline{x}(n)$. [avec un produit scalaire défini par $\langle x, y \rangle = E[x y^*]$ sur l'espace des signaux aléatoires de puissance moyenne finie].

Le gradient s'annule pour

$$E[\underline{x}^{(n)*} x^{(n)}] = E[\underline{x}^{(n)*} \underline{x}^{(n)T}] \underline{\alpha}$$

Où l'on reconnaît la matrice de corrélation définie par

$$\underline{R}_x = E[\underline{x}^{(n)*} \underline{x}^{(n)T}] = \begin{bmatrix} R(0) & R(1) & \dots & R(p-1) \\ \vdots & \vdots & \ddots & \vdots \\ R(1-p) & \dots & \dots & R(0) \end{bmatrix}$$

et le vecteur de corrélation

$$\underline{r}_x = E[\underline{x}^{(n)*} x^{(n)}]^T = [R(1) \dots R(p)]^T$$

Ceci fournit finalement

$$\underline{R}_x \underline{\alpha}^* = \underline{r}_x^*$$

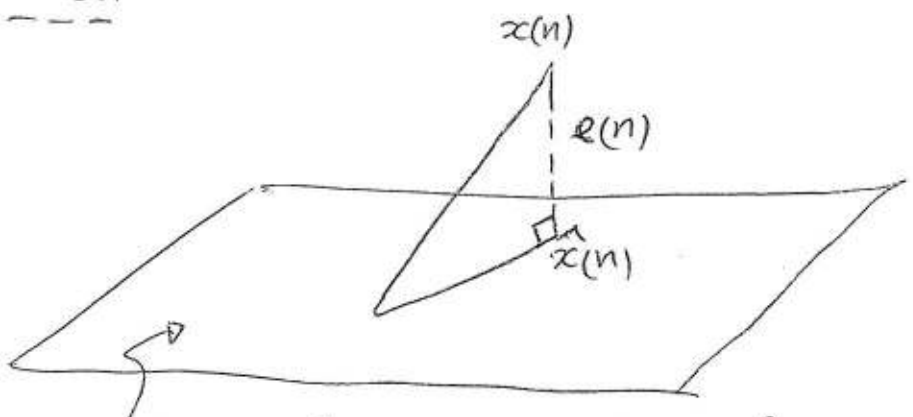
NB - On note parfois

$\underline{R}_x = E[\underline{x}^{(n)*} \underline{x}^{(n)T}]$, ce qui conduirait à

$$\underline{R}_x \underline{\alpha} = \underline{r}_x$$

Cette relation est appelée équation de Yule Walker, ou équation normale, et permet de déterminer les coefficients de prédiction.

On a noté qu'à l'optimum, l'erreur $e(n)$ est orthogonale aux données $x(n)$. Le résultat est appelé théorème de projection



hyperplan engendré par $\{x^{(n-1)}, \dots, x^{(n-p)}\}$.

Signaux Auto-régressifs

On appelle signal autorégressif d'ordre p un signal aléatoire défini par l'équation aux différences

$$x(n) = \sum_{i=1}^p a_i x(n-i) + b(n),$$

où $b(n)$ est un bruit blanc de variance σ_b^2 .
Il s'agit de la sortie d'un filtre tout pôles, de fonction de transfert

$$H(z) = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}},$$

excité par un bruit blanc.

Bien entendu, la meilleure prédiction à un pas à l'aide d'un passé de longueur p est donnée par

$$\hat{x}(n) = \sum_{i=1}^p d_i x(n-i) = \sum_{i=1}^p a_i x(n-i)$$

et l'on a alors coincidence des coefficients de prédiction et des coefficients AR.

En posant $\underline{a} = [a_1 \ a_2 \ \dots \ a_p]^T$

$$\underline{x}(n) = [x(n-1) \ \dots \ x(n-p)]$$

on obtient

$$\begin{aligned} x(n) &= \underline{a}^T \underline{x}(n) + b(n) \\ &= \underline{x}(n)^T \underline{a} + b(n) \end{aligned}$$

En multipliant à gauche par $\underline{x}(n)^*$ et en prenant l'espérance, il vient:

$$E[\underline{x}(n)^* \underline{x}(n)^T] \underline{a} = E[\underline{x}(n)^* x(n)] - \underbrace{E[b(n) \underline{x}(n)]}_{=0}$$

où l'on reconnaît exactement la même relation que dans le cas de la prédiction linéaire, soit

$$\underline{R}_x \underline{a}^* = \underline{R}_x^*$$

avec \underline{R}_x de taille $P \times P$.

Il est ainsi possible d'identifier aisément les coefficients du modèle AR, par simple résolution d'un système linéaire. [notons que l'équation aux différences générique est causale et que le filtre sous-jacent $H(z)$ est stable. Les pôles correspondant sont donc de module < 1 : la modélisation AR ne peut identifier correctement qu'un système causal et stable].

Notons que l'on a également

$$E[x(n) b(n)^*] = E[b(n)^* \underline{x}(n)^T] \underline{a} + E[b(n)^* b(n)]$$

$$= 0 + \sigma^2$$

car $E[b(n)^* x(n-k)] = 0$ pour $k > 0$,
par causalité.

On peut alors augmenter l'ordre du système afin d'identifier les coefficients AR et la variance du bruit générateur σ_b^2 :

Pour cela, on réécrit l'éq. aux différences sous la forme

$$\begin{bmatrix} x(n) & \underline{x}(n)^T \end{bmatrix} \begin{bmatrix} 1 \\ -\underline{a} \end{bmatrix} = b(n)$$

En notant

$$\underline{x}(n+1) = \begin{bmatrix} x(n) & \underline{x}(n)^T \end{bmatrix}^T$$

et en multipliant à gauche par $\underline{x}(n+1)^*$ et en prenant l'espérance, on obtient

$$E \left[\underline{x}(n+1)^* \underline{x}(n+1)^T \right] \begin{bmatrix} 1 \\ -a \end{bmatrix} = E \left[b(n) \underline{x}(n+1)^* \right]$$

soit

$$R_{x,p+1} \cdot \begin{bmatrix} 1 \\ -a \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Ou encore

$$R_{x,p+1} \cdot \begin{bmatrix} 1/\sigma^2 \\ -a/\sigma^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

À partir de cette relation, on peut identifier à la fois les coefficients AR et la variance du bruit générateur σ^2 .

Notons que la matrice de corrélation d'ordre $(p+1)$, $R_{x,p+1}$ peut être partitionnée selon

$$\left[\begin{array}{c|c} R_{xx}(0) & \underline{R}_{xx,p}^+ \\ \hline \underline{R}_{xx,p} & R_{xx,p} \end{array} \right] \begin{bmatrix} 1 \\ -a \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

avec $\underline{R}_{xx,p}$ la matrice d'ordre p et $\underline{R}_{x,p}$ le vecteur d'intercorrélation. On retrouve alors l'équation normale d'ordre p en développant: $\underline{R}_x - \underline{R}_{x,p} \underline{a} = 0$.

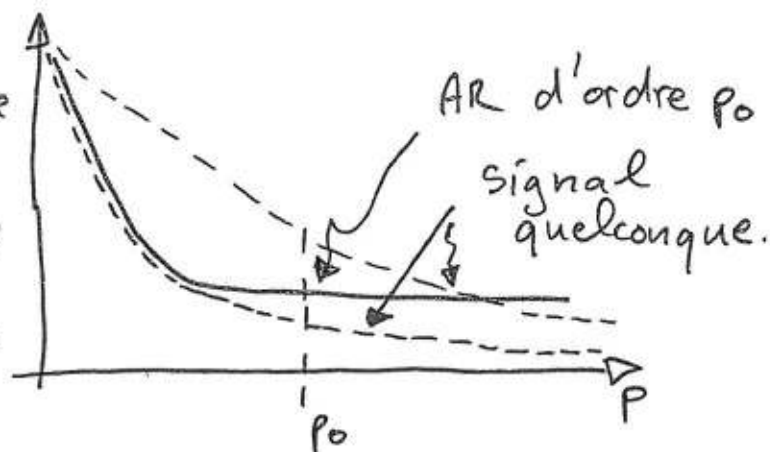
Dans le cas d'un AR(p) et d'une prédiction avec un passé de longueur p , les coefficients de prédiction \hat{a} s'identifient aux paramètres AR a , et l'erreur de prédiction $e(n)$ se confond avec le bruit générateur du modèle AR, $u(n)$. L'erreur quadratique vaut alors

$$\sigma^2 = E[|u(n)|^2] = \sigma^2$$

et $e(n)$ est un bruit blanc.

Il est important de noter que si la prédiction est réalisée avec un passé de longueur inférieure à p , $\hat{x}(n)$ sera prédit de manière « incomplète », et l'erreur de prédiction correspondante ne sera pas blanche et aura une puissance $> \sigma^2$. Lorsque l'ordre de prédiction atteint ou dépasse p , l'erreur de prédiction est un bruit blanc totalement imprédictible et sa variance vaut σ^2 . La courbe suivante représente la puissance de l'erreur lorsque p croît.

Pour un signal quelconque (non AR), l'erreur de prédiction décroît avec p , le système tentant de « blanchir » le résiduel.



Choix de l'ordre - Critères de sélection

Le choix de l'ordre du processus, cet ordre n'est bien entendu pas connu a priori, peut-être très important. En effet, si on sous-estime l'ordre de prédiction, le modèle ne sera pas adapté, et la densité spectrale estimée ne pourra faire apparaître les détails pertinents : elle sera trop douce (ou « molle »). Par contre, si on surestime l'ordre, l'augmentation du nombre de paramètres du modèle entraîne, en pratique, une dégradation de la tenue statistique, bien que le modèle soit d'ordre suffisant pour bien représenter la dsp.

Nous avons vu que la puissance de l'erreur de prédiction décroît lorsque l'ordre du modèle augmente, en atteignant éventuellement un palier ou elle reste constante, dans le cas d'un « vrai » signal AR.

Un compromis doit être trouvé entre une augmentation de l'ordre qui permet de mieux modéliser le processus et une limitation du nombre de paramètres du modèle pour préserver la variance.

De nombreux critères de sélection de l'ordre ont été proposés. Ces critères sont en général composés d'une fonction de l'erreur de prédiction, décroissante avec l'ordre, et d'une fonction croissante de l'ordre. Un ordre optimum est alors sélectionné en recherchant le minimum du critère.

$$\text{FPE}(p) = \frac{N+p}{N-p} \sigma_p^2 \quad \text{Final Prediction Error}$$

$$\text{AIC}(p) = N \log_e \sigma_p^2 + 2p \quad \text{Akaike Information Criterion}$$

$$\text{MDL}(p) = \log_e \sigma_p^2 + \frac{p+1}{N} \log_e(N) \quad \text{Minimum Description Length}$$

$$\hat{\epsilon}_{min} = \begin{bmatrix} \frac{\lambda}{2} & 0 \end{bmatrix} \underline{w}$$

$$\hat{\epsilon}_{min} = \frac{\lambda}{2}$$

En se souvenant, que l'erreur quadratique minimale sera une estimée de la puissance de l'erreur de prédiction (au facteur $1/(b-a)$ près), il vient

$$\underline{w} = (\underline{X}_k^+ \underline{X}_k)^{-1} \begin{bmatrix} (b-a) \hat{\sigma}_p^2 \\ \vdots \\ 0 \end{bmatrix}$$

Cette relation a exactement la même structure que les équations normales, en remplaçant R_{xx} par $(\underline{X}_k^+ \underline{X}_k)$. En fait, on s'aperçoit en développant $(\underline{X}_k^+ \underline{X}_k)$ que ce produit constitue une estimée de la matrice de corrélation. En particulier, dans le cas ②, $\underline{X}_k^+ \underline{X}_k$ a une structure de Toeplitz, et l'algorithme de Levinson peut-être utilisé pour la résolution. On parle dans ce cas de méthode de corrélation. Dans les hypothèses ①, ③, ④, la matrice obtenue n'a pas la structure Toeplitz, mais des algorithmes efficaces ont été développés. Notons que la méthode correspondant à l'hypothèse ① est appelée méthode de covariance. Cette méthode est utilisée lorsqu'on veut limiter les « effets de bord », c'est-à-dire en particulier lorsque le nombre de données disponibles est faible.

Estimation de la fonction de corrélation.

On dispose d'un signal à temps discret sur N points, $n=0 \dots N-1$. On utilise éventuellement une fenêtre de pondération $g_N(n)$ sur les données (longueur N) et l'on pose $x_N(n) = g_N(n) x(n)$

Si le signal sous-jacent est stationnaire ergodique, on sait que

$$R_{xx}(k) = E[X(n) X^*(n-k)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{[N]} X(n) X^*(n-k)$$

Dans notre cas, on ne dispose que d'un nombre fini d'échantillons, et la quantité

$$\hat{R}_{xx}(k) = \frac{1}{N} \sum X(n) X^*(n-k)$$

ne nous donnera qu'une valeur approchée, estimée de la corrélation. Dans la mesure où $\hat{R}_x(k)$ est construit à partir d'une somme finie de variables aléatoires, $\hat{R}(k)$ sera elle-même une fonction aléatoire dont il faudrait étudier les propriétés.

De manière générale, on pose

$$\hat{R}_{xx}(k) = \alpha_{N,k} \sum_{n=-\infty}^{+\infty} X_N(n) X_N^*(n-k)$$

où $\alpha_{N,k}$ est un coefficient qui peut dépendre du retard k , et où l'on a pris des portées des $-\infty$ à $+\infty$, puisque la limitation temporelle est comprise dans la fenêtre $g_N(n)$ appliquée sur les données.

Puisque $\hat{R}_x(k)$ est aléatoire, il est naturel de s'intéresser à son comportement moyen:

$$\begin{aligned} E[\hat{R}_x(k)] &= E\left[\alpha_{N,k} \sum_n g_N(n) X(n) g_N(n-k) X^*(n-k)\right] \\ &= \alpha_{N,k} \sum_{n=-\infty}^{+\infty} g_N(n) g_N^*(n-k) E[X(n) X^*(n-k)] \\ &= \alpha_{N,k} \cdot w_N(k) \cdot R_{xx}(k) \end{aligned}$$

avec $w_N(k) = \sum_{n=-\infty}^{+\infty} g_N(n) g_N^*(n-k) = (g * g^*)(k)$.

L'estimateur est dit non-biaisé si on prend $\alpha_{N,k} = \frac{1}{w_N(k)}$ (pour $w_N(k) \neq 0$) évidemment.

Si l'on choisit, même implicitement, la fenêtre naturelle $\pi_N(n) = 1$ pour $n=0 \dots N-1$ et $= 0$ sinon, on obtient $w_N(k) = N \cdot \text{Tri}_N(k)$ un triangle d'amplitude N et de base de longueur $2N-1$ $w_N(k) = N - |k|$. On utilise principalement

$$\begin{aligned} \hat{R}_{xx}^{NB}(k) &= \frac{1}{N-|k|} \sum_{n=0}^{N-1-|k|} X(n) X^*(n-k) \quad k < 0 \\ &= \frac{1}{N-|k|} \sum_{n=k}^{N-1} X(n-1) X^*(n-k) \quad k \geq 0 \end{aligned}$$

Qui est l'estimateur dit non biaisé [mais qui est biaisé : en effet, on ne peut calculer $\hat{R}_x(k)$ que pour $|k| \leq (N-1)$, et

$$E[\hat{R}_{xx}^{(NB)}(k)] = R_{xx}(k) \cdot \pi_{2N-1}(k)$$

$$\begin{aligned}\hat{R}_{xx}^B(k) &= \frac{1}{N} \sum_{n=0}^{N-1-|k|} X(n) X^*(n-k) \quad k < 0 \\ &= \hat{R}_{xx}^*(k) \quad [\text{symétrie hermitienne}]\end{aligned}$$

est très employé, pour 2 faisceaux : son comportement en terme de variance, et le fait que sa TF reste positive.

On notera que l'on a simplement

$$\hat{R}_x^{(B)}(k) = \text{Tri}_N(k) \cdot \hat{R}^{(NB)}(k).$$

Calcul de la variance:

Pour la fenêtre naturelle $g_N(n) = \pi_N(n)$, on a

$$E[\hat{R}_{xx}(k)] = \alpha_{N,k} (N-|k|) R_{xx}(k).$$

La variance est donnée par

$$\text{Var}[\hat{R}_{xx}(k)] = E[|\hat{R}_{xx}(k)|^2] - |E[\hat{R}_{xx}(k)]|^2$$

Le moment d'ordre 2 s'exprime comme

$$\begin{aligned}E[|\hat{R}_{xx}(k)|^2] &= E\left[|\alpha_{k,n}|^2 \sum_n X(n) X^*(n-k) \sum_m X(m) X^*(m-k)\right] \\ &= |\alpha_{k,n}|^2 \sum_n \sum_m E[X(n) X^*(n-k) X(m) X^*(m-k)]\end{aligned}$$

Et le calcul s'arrête là, sans hypothèse supplémentaire. Pour aller plus avant, supposons donc que $X(n)$ est gaussien réel.

Pour une variable gaussienne centrée, on a

$$E[x_1 x_2 x_3 x_4] = E[x_1 x_2] E[x_3 x_4] + E[x_1 x_3] E[x_2 x_4] + E[x_1 x_4] E[x_2 x_3] -$$

Ce qui fournit ici

$$E[x^{(n)} x^{(n-k)} x^{(m)} x^{(m-k)}] = R(k)^2 + R(n-m)^2 + R(n-m+k) R(n-m-k)$$

Et

$$\text{Var} [\hat{R}_x(k)] = \alpha_{N,k}^2 \sum_n \sum_m R(k)^2 + R(n-m)^2 + R(n-m+k) R(n-m-k) - \alpha_{N,k}^2 (N-k)^2 R(k)^2$$

soit

$$\text{Var} [\hat{R}_x(k)] = \alpha_{N,k}^2 \sum_n \sum_m R(n-m)^2 + R(n-m+k) R(n-m-k)$$

En posant $l = n-m$,

$$\text{Var} [\hat{R}_x(k)] \approx \alpha_{N,k}^2 \sum_l (R(l)^2 + R(l+k) R(l-k)) w(l)$$

où $w(l)$ représente la multiplicité du terme $R(l)$: on obtient par exemple $R(0)$ en prenant $n=m$, pour $n=0 \dots N-1$, soit N termes, tandis que $R(N-1)$ n'est obtenu que pour $n=N-1$ et $m=0$, un seul terme - Finalement, $w(l)$ est une fonction triangle : $w(l) = N - |l|$, et

$$\text{Var} [\hat{R}_x(k)] = N \cdot \alpha_{N,k}^2 \sum_{l=-(N-1)}^{N-1} \left(1 - \frac{|l|}{N}\right) \left[R_x(l)^2 + R_x(l+k) R_x(l-k) \right]$$

Pour l'estimateur non biaisé, on

$$\alpha_{N,k} = \frac{1}{N-|k|} \rightarrow \text{Var} [\hat{R}_x(k)] = \frac{N}{(N-|k|)^2} \sum_e$$

$$\sim \frac{1}{N-|k|} \sum_e$$

Pour $k \rightarrow N$, le terme $\frac{1}{N-k}$ \nearrow et tend vers l' ∞ alors que la \sum_e reste finie : la variance diverge ce qui signifie que l'estimateur « explose » sur les bords de l'intervalle sur k .

Pour l'estimateur biaisé, $\alpha_{N,k} = \frac{1}{N}$ et $\text{Var} \hat{R}_x(k) = \frac{1}{N} \sum_e$, variance qui est toujours plus faible que celle de l'estimateur non biaisé. Les 2 estimateurs sont convergents dans le sens où $\text{Var} \rightarrow 0$ quand $N \rightarrow \infty$.

Transformation de Fourier de l'estimateur de corrélation

Par le théorème de Wiener-Kintchine, on sait que la densité spectrale de puissance est la TF de la fonction de corrélation. La TF d'un estimateur de la corrélation nous fournit donc un estimateur de la densité spectrale de puissance.

On notera $\hat{S}_{xx}^{(NB)}(\omega) = \text{TF} \{ \hat{R}_{xx}^{(NB)} \}$ et $\hat{S}_{xx}^{(B)}(\omega) = \text{TF} \{ \hat{R}_{xx}^{(B)} \}$

Il est aisé de vérifier que

$$E[\hat{S}_{xx}^{(-)}(\omega)] = E[\text{TF} \{ \hat{R}_{xx}^{(-)} \}] = \text{TF} \{ E[\hat{R}_{xx}^{(-)}] \}$$

cas NB :

$$E[\hat{S}_{xx}^{NB}(\omega)] = \text{TF} \{ \frac{1}{2N+1} R_{xx}(k) \} = \text{sinc} * S_{xx}(\omega)$$

cas B :

$$E[\hat{S}_{xx}^B(\omega)] = \text{TF} \{ N \text{Tri}_N(k) R_{xx}(k) \} = (\text{sinc})^2 * S_{xx}(\omega)$$

Le premier estimateur, à cause de la convolution par un (sinc), peut donc conduire à une dsp présentant éventuellement des portions négatives, ce qui est assez peu physique - Le second, au contraire conduit à une dsp toujours positive, ce qui est plus raisonnable - On devra donc éviter le premier estimateur.

Dans le second cas, le résultat exact est

$$E[\hat{S}_{xx}^{(B)}(f)] = \left(\frac{\sin(\pi f N)}{N \cdot \sin \pi f} \right)^2 * S_{xx}(f).$$

En outre, il est possible de calculer directement $\hat{S}_{xx}^{(B)}(f)$ à partir des données:

TF de l'estimateur biaisé - Périodogramme

$$\begin{aligned} \hat{S}_{xx}^{(B)}(f) &= \text{TF} \left\{ \hat{R}_{xx}^{(B)}(k) \right\} = \sum_k \frac{1}{N} \sum_n X_N(n) X_N^*(n-k) e^{-j2\pi f k} \\ &= \frac{1}{N} \sum_k \sum_n X_N(n) e^{-j2\pi f n} X_N^*(n-k) e^{+j2\pi f(n-k)} \\ &= \frac{1}{N} \sum_k \underbrace{X_N(n) e^{-j2\pi f n}}_{X_N(f)} \sum_{v=n-k} X_N^*(v) e^{j2\pi f v} \\ &= \frac{1}{N} X_N(f) \cdot X_N(f)^* \\ &= \frac{1}{N} |X_N(f)|^2 \end{aligned}$$

où $X_N(n)$ représente les données sur N points. Il est donc possible de calculer $\hat{S}_x^{(B)}(f)$ comme le module carré de la TF des données, sans passer par l'estimation d'une corrélation. La quantité $\hat{S}_{xx}^{(B)}(f)$ est appelée périodogramme.

Le périodogramme $\hat{S}_{xx}^{(B)}(f) = \text{TF} \left\{ \hat{R}_{xx}^{(B)} \right\} = \frac{1}{N} |X_N(f)|^2$ 40
 est donc toujours positif, calculable directement à partir de la TF des données.

Du fait de la convolution par un (sinc)², le périodogramme présente une limitation en résolution en $\frac{1}{N}$, la résolution de Fourier -

Corrélogramme.

On a déjà noté que $\hat{R}_{xx}^{(B)}(k) = \left(1 - \frac{|k|}{N}\right) \hat{R}_x^{(NB)}(k)$.

On peut généraliser ceci à une fenêtre quelconque $w(k)$ de longueur N :

$$\hat{R}_{xx}^{(\text{cor})}(k) = w(k) \cdot \hat{R}_{xx}^{(NB)}(k)$$

On appelle corrélogramme la dsp estimée par TF de $\hat{R}_{xx}^{(\text{cor})}(k)$:

$$\hat{S}_{xx}^{(\text{cor})} = \sum_{k=-(N-1)}^{N-1} w(k) \hat{R}_{xx}^{(NB)}(k) e^{-j2\pi f k}$$

Biais du corrélogramme:

$$\text{On a } \hat{S}_{xx}(f) = \sum_{k=-(N-1)}^{N-1} w(k) \hat{R}_{xx}^{NB}(k) e^{-j2\pi f k}$$

Comme l'estimateur de la corrélation est non biaisé, $E[\hat{R}_{xx}^{NB}(k)] = R_{xx}(k)$, la moyenne du corrélogramme vaut

$$E[\hat{S}_{xx}(f)] = \sum_{k=-(N-1)}^{N-1} w(k) R_{xx}(k) e^{-j2\pi f k}$$

En utilisant pour $w(k)$

- la fenêtre rectangulaire sur $[-(N-1), N-1]$ dans le cas de l'estimateur NB,
- la fenêtre triangulaire sur $[-(N-1), N-1]$ dans le cas de l'estimateur B,
- une fenêtre générale,

on peut écrire

$$E[\hat{S}_{xx}(f)] = \sum_{k=-\infty}^{+\infty} w(k) R_{xx}(k) e^{-j2\pi f k}$$

Il s'agit donc simplement de la TF d'un produit de deux séquences et

$$E[\hat{S}_{xx}(f)] = W(f) * S_{xx}(f)$$

$$\text{où } W(f) = \frac{\sin(\pi f N)}{\sin(\pi f)}$$

pour la fenêtre rectangulaire sur $[-(N-1); N-1]$

$$= \frac{1}{N} \left(\frac{\sin(\pi f N)}{\sin(\pi f)} \right)^2$$

pour la fenêtre triangulaire

D'une manière générale, on constate donc que le corrélogramme est biaisé, et est plus « étalé » que la vraie densité spectrale de puissance (du

fait de la convolution). Par voie de conséquence, le corrélogramme est limité en résolution (la résolution reflète la capacité à séparer deux raies pures).

Les fenêtres de pondération étant de largeur maximale N (pour des méthodes mises en œuvre à partir de signaux connus sur N points), la résolution obtenue est inversement proportionnelle à N .

Ceci est caractéristique des méthodes d'analyse spectrale « classique », ou méthodes de Fourier, où la résolution est limitée par le nombre de points de données.

On appelle résolution de Fourier la quantité $\frac{1}{N}$

L'étude de la moyenne du corrélogramme nous montre donc que celui-ci permet d'approcher la densité spectrale de puissance, avec un défaut apporté par la limitation en résolution liée à l'utilisation de fenêtres de pondération. Notons bien que ces fenêtres de pondération sont de toute manière présentes (fenêtre naturelle), du fait de la limitation temporelle des signaux. Nous examinerons plus loin les caractéristiques des principales fenêtres de pondération et les critères de choix.

Ceci étant posé, il convient de se souvenir du fait que les dsp estimées, comme transformées de signaux aléatoires, sont elles mêmes aléatoires, et il apparaît alors des « fluctuations » autour de la valeur moyenne. Une manière de quantifier l'importance de ces fluctuations est de s'intéresser à la variance des estimateurs..

Nous débiterons par l'étude de la variance du périodogramme. Nous avons introduit celui-ci comme le corrélogramme particulier associé à l'estimateur biaisé de la séquence de corrélation :

$$S_{xx}^{PER}(\delta) = TF \left\{ \hat{R}_{xx}^B(k) \right\} = \frac{1}{N} |X_N(\delta)|^2.$$

Le moment d'ordre 2 du périodogramme vaut alors :

$$E[S_{xx}^{\text{PER}}(f_1) S_{xx}^{\text{PER}}(f_2)] = \frac{1}{N^2} E[|X_N(f_1)|^2 |X_N(f_2)|^2]$$

avec $X_N(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi n f}$.

Dans ces conditions,

$$E[S_{xx}^{\text{PER}}(f_1) S_{xx}^{\text{PER}}(f_2)] = \frac{1}{N^2} E\left[\sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(k) x(l) x(m) x(n) \times e^{-j2\pi[(k-l)f_1 + (m-n)f_2]}\right].$$

$$= \frac{1}{N^2} \sum_k \sum_l \sum_m \sum_n E[x(k) x(l) x(m) x(n)] e^{-j2\pi[(k-l)f_1 + (m-n)f_2]}$$

Pour aller plus loin, nous allons faire deux hypothèses :

- o x est un processus gaussien de variance σ^2
- o x est un bruit blanc, i.e. $R_{xx}(k) = \sigma^2 \delta(k)$

Si x est gaussien, alors on peut exprimer le moment d'ordre 4 comme un produit des moments d'ordre 2 :

$$\begin{aligned} E[x(k) x(l) x(m) x(n)] &= E[x(k) x(l)] E[x(m) x(n)] \\ &\quad + E[x(k) x(m)] E[x(l) x(n)] + E[x(k) x(n)] E[x(l) x(m)] \\ &= R_{xx}(k-l) R_{xx}(m-n) + R_{xx}(k-m) R_{xx}(l-n) \\ &\quad + R_{xx}(k-n) R_{xx}(l-m) \end{aligned}$$

Enfin, puisque x est blanc,

$$E[x(k) x(l) x(m) x(n)] = \sigma^4 \left[\delta(k-l) \delta(m-n) + \delta(k-m) \delta(l-n) + \delta(k-n) \delta(l-m) \right]$$

Cette expression vaut

$$\begin{aligned} &\sigma^4 \quad \text{si } \begin{cases} k=l & m=n \\ k=m & l=n \\ k=n & l=m \end{cases} \\ &3\sigma^4 \quad \text{si } k=l=m=n \\ &0 \quad \text{sinon} \end{aligned}$$

On en déduit que

$$\begin{aligned}
 E[S_{xx}^{\text{PER}}(f_1) S_{xx}^{\text{PER}}(f_2)] &= \frac{1}{N^2} \sum_{k=l} \sum_{m=n} \sigma^4 \\
 &+ \frac{1}{N^2} \sum_{k=m}^{N-1} \sum_{l=n=0}^{N-1} \sigma^4 e^{-j2\pi[(k-l)f_1 + (m-n)f_2]} \\
 &+ \frac{1}{N^2} \sum_{k=n=0}^{N-1} \sum_{m=l=0}^{N-1} \sigma^4 e^{-j2\pi[(k-l)f_1 + (m-n)f_2]}
 \end{aligned}$$

En posant $u = k-l$, $v = m-n$
 En notant que $(k-l) = \pm(m-n)$ et que la première double somme comprend N^2 termes,

$$\begin{aligned}
 &= \sigma^4 + \sigma^4 \frac{1}{N^2} \sum_k e^{-j2\pi k(f_1+f_2)} \\
 &+ \sigma^4 \frac{1}{N^2} \sum_k e^{-j2\pi k(f_1-f_2)}
 \end{aligned}$$

$$\text{soit } E[S_{xx}^{\text{PER}}(f_1) S_{xx}^{\text{PER}}(f_2)] = \sigma^4 \left[1 + \left(\frac{\sin \pi N (f_1+f_2)}{N \sin \pi (f_1+f_2)} \right)^2 + \left(\frac{\sin \pi N (f_1-f_2)}{N \sin \pi (f_1-f_2)} \right)^2 \right]$$

La covariance du périodogramme vaut alors

$$E[S_{xx}^{\text{PER}}(f_1) S_{xx}^{\text{PER}}(f_2)] - E[S_{xx}^{\text{PER}}(f_1)] E[S_{xx}^{\text{PER}}(f_2)]$$

Pour un bruit blanc, $S_{xx}(f) = \sigma^2$, et le périodogramme est non-biaisé - La covariance vaut donc

$$\text{Cov}[S_{xx}^{\text{PER}}(f_1), S_{xx}^{\text{PER}}(f_2)] = \sigma^4 \left[\left(\frac{\sin \pi N (f_1+f_2)}{N \sin \pi (f_1+f_2)} \right)^2 + \left(\frac{\sin \pi N (f_1-f_2)}{N \sin \pi (f_1-f_2)} \right)^2 \right]$$

Pour N grand,

$$\left(\frac{\sin(\pi f_0 N)}{N \sin \pi f_0} \right)^2 \rightarrow \mathbb{1}(f_0)$$

Fonction indicatrice
 avec $\begin{cases} \mathbb{1}(f) = 0 \text{ pour } f \neq 0 \\ \mathbb{1}(0) = 1 \end{cases}$

$$\text{Cov}[S_{xx}^{\text{PER}}(f_1), S_{xx}^{\text{PER}}(f_2)] = S_{xx}(f_1) S_{xx}(f_2) [\mathbb{1}(f_1+f_2) + \mathbb{1}(f_1-f_2)]$$

Par conséquent pour $f_1 \neq \pm f_2$ la covariance est nulle et les composantes spectrales sont (asymptotiquement) décorréelées.

Enfin, la variance à une fréquence donnée f vaut

$$\text{Var} [S_{xx}^{\text{PER}}(f)] = \text{Cov} [S_{xx}^{\text{PER}}(f), S_{xx}^{\text{PER}}(f)]$$

$$\text{Var} [S_{xx}^{\text{PER}}(f)] = S_{xx}(f)^2 \left[1 + \left(\frac{\sin(2\pi f N)}{N \sin 2\pi f} \right)^2 \right].$$

On constate donc, et ce résultat est général, que la variance du périodogramme est toujours supérieure au carré de la densité spectrale recherchée. Il est important de noter que ceci est vrai quelque soit le nombre de points disponible : lorsque $N \rightarrow \infty$, on a simplement

$$\begin{cases} \text{Var} [S_{xx}^{\text{PER}}(0)] = 2 S_{xx}(0)^2 \\ \text{Var} [S_{xx}^{\text{PER}}(f)] = S_{xx}(f)^2 \text{ pour } f \neq 0 \end{cases}$$

L'estimateur du périodogramme n'est pas consistant : sa variance reste finie lorsque le nombre de points tend vers l'infini - L'estimateur converge en moyenne, puisque

$$E [S_{xx}^{\text{PER}}(f)] = S_{xx}(f) \text{ quand } N \rightarrow \infty,$$

mais il ne converge pas en moyenne quadratique puisque

$$E [(S_{xx}^{\text{PER}}(f) - S_{xx}(f))^2] = S_{xx}(f)^2 \left[1 + \left(\frac{\sin 2\pi f N}{N \sin 2\pi f} \right)^2 \right].$$

L'estimateur du périodogramme n'est donc pas un bon estimateur, dans la mesure où l'écart-type sur la valeur à estimer est supérieur à la valeur recherchée... Il est donc indispensable de chercher à améliorer la « qualité » de cet estimateur.

Rappelons que nous avons défini le périodogramme comme le corrélogramme particulier associé à l'estimateur biaisé de la corrélation. Il convient donc maintenant de s'intéresser aux corrélogrammes utilisant des fenêtres de pondération plus générales.

Le corrélogramme est défini par

$$\hat{S}_{xx}^{(COR)}(f) = \sum_{-\infty}^{+\infty} w(k) \hat{R}_{xx}^{(B)}(k) e^{-j2\pi f k}$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} W(f-u) \hat{S}_{xx}^{PER}(f) df = W(f) * \hat{S}_{xx}^{PER}(f)$$

Puisque $E[\hat{S}_{xx}^{PER}(f)] = \left(\frac{\sin \pi f N}{N \sin \pi f}\right)^2 * S_{xx}(f)$, la moyenne du corrélogramme est

$$E[\hat{S}_{xx}^{COR}(f)] = W(f) * \left(\frac{\sin \pi f N}{N \sin \pi f}\right)^2 * S_{xx}(f)$$

Afin de ne pas faire apparaître de puissance négative, il faut que $W(f) \geq 0$ (et que $W(f)$ soit également réelle, ce qui entraîne $w(k)$ à symétrie hermitienne).

L'introduction d'une fenêtre de pondération $w(k)$ sur la corrélation correspond donc, dans le domaine fréquentiel, à effectuer un lissage du périodogramme.

Calcul de la variance.

Il est facile d'imaginer que ce lissage, c'est-à-dire un moyennage pondéré des composantes spectrales adjacentes décorréliées, entraîne une réduction de la variance.

$$\text{Var}[\hat{S}_{xx}^{COR}(f)] = E\left[\left| P(f) * \left(\hat{S}_{xx}^{PER}(f) - E[\hat{S}_{xx}^{PER}(f)] \right) \right|^2 \right]$$

avec $P(f) = W(f) * \left(\frac{\sin \pi f N}{N \sin \pi f}\right)^2$

$$\text{soit } \text{Var}[\hat{S}_{xx}^{COR}(f)] = E\left[\int P(f-u) \Delta S_{xx}^{PER}(u) du \cdot \left(\int P(f-v) \Delta S_{xx}^{PER}(v) dv \right)^* \right]$$

$$= \iint P(f-u) P(f-v)^* E\left[\Delta S_{xx}^{PER}(u) \cdot \Delta S_{xx}^{PER}(v) \right] dudv$$

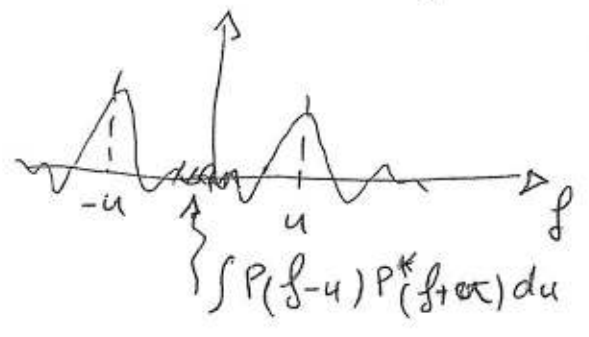
$$= \iint P(f-u) P(f-v)^* \text{COV}\left[\hat{S}_{xx}^{PER}(u), \hat{S}_{xx}^{PER}(v) \right] dudv$$

Or $Cov \left[\hat{S}_{xx}^{PER}(u), \hat{S}_{xx}^{PER}(v) \right] \sim \frac{S_{xx}(u) S_{xx}(v)}{N} \left[\mathbb{1}(u+v) + \mathbb{1}(u-v) \right]$
 et la variance devient donc $\sim \frac{S_{xx}(u) S_{xx}(v)}{N} \left[\delta(u+v) + \delta(u-v) \right]$

$$Var \left[\hat{S}_{xx}^{COR}(f) \right] \sim \int P(f-u) P^*(f+u) S_{xx}^2(u) + |P(f-u)|^2 S_{xx}^2(u) du$$

Le premier terme est petit devant le second si f n'est pas trop proche de $f=0$ et $\pm \frac{1}{2}$.

Par ailleurs, si la dsp $S_{xx}(u)$ est à variations lentes sur la bande du lobe principal de $P(f)$, alors



$$\begin{aligned} Var \left[\hat{S}_{xx}^{COR}(f) \right] &\sim \frac{1}{N} S_{xx}(f)^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} |P(f-u)|^2 du \\ &= \frac{1}{N} S_{xx}(f)^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} |P(u)|^2 du \\ &= S_{xx}(f)^2 \frac{1}{N} \sum_n P(n)^2 \quad [Parseval] \end{aligned}$$

En analyse spectrale classique, on distingue et on présente souvent séparément « la méthode du corrélogramme » et les méthodes du périodogramme.

En fait, nous avons déjà vu apparaître le périodogramme comme corrélogramme particulier associé à $\hat{R}_{xx}^B(k)$, et nous avons interprété les méthodes du corrélogramme comme conduisant à un lissage du périodogramme.

Il nous reste cependant à présenter la méthode du périodogramme moyenné, avant d'ajouter quelques commentaires sur le lissage du périodogramme.

Périodogramme moyenné :

On a vu que le périodogramme est asymptotiquement non biaisé :

$$E \left[S_{xx}^{PER}(f) \right] = T_N(f) * S_{xx}(f)$$

$$\rightarrow S_{xx}(f) \text{ lorsque } N \rightarrow +\infty$$

La variance du périodogramme est alors liée à l'absence d'opérateur de moyenne statistique $E[-]$. Une manière d'approcher l'opérateur espérance mathématique est d'utiliser un moyennage de périodogrammes calculés sur des réalisations distinctes du signal. Évidemment, en pratique on dispose rarement de plusieurs et a fortiori de nombreuses réalisations d'un signal aléatoire. Une idée est alors de découper une seule réalisation (les données disponibles) en tranches sur lesquelles on calcule un ensemble de périodogrammes qui sont ensuite moyennés.

On découpe donc un signal de N points en L segments de K points.

$$x_k^{(i)}(n) = x(i.L + n) \quad \text{pour } i \in [0, L-1] \\ n \in [0, K-1]$$

Un périodogramme élémentaire s'écrit alors

$$\begin{aligned}\hat{S}_{xx}^{PER(i)}(f) &= \frac{1}{K} \left| X_K^{(i)}(f) \right|^2 \\ &= \frac{1}{K} \left| \sum_{n=0}^{K-1} x^{(i)}(n) e^{-j2\pi n f} \right|^2,\end{aligned}$$

et l'estimateur du périodogramme moyenné s'en déduit simplement:

$$\hat{S}_{xx}^{MOY}(f) = \frac{1}{L} \sum_{i=0}^{L-1} \hat{S}_{xx}^{PER(i)}(f).$$

Si les segments sont indépendants (ou décorrelés dans le cas gaussien), alors la variance vaut simplement

$$\begin{aligned}\text{Var}[\hat{S}_{xx}^{MOY}(f)] &= \frac{1}{L} \text{Var}[\hat{S}_{xx}^{PER}(f)] \\ &= \frac{1}{L} S_{xx}(f)^2 \left[1 + \left(\frac{\sin 2\pi f K}{K \sin 2\pi f} \right)^2 \right].\end{aligned}$$

La variance est donc diminuée par un facteur $1/L$. Notons que l'hypothèse d'indépendance ou de décorrélation des différents segments n'est pas particulièrement raisonnable: ceci signifierait que les échantillons sont décorrelés d'une tranche à l'autre, c'est-à-dire finalement que le processus est blanc. En pratique, la variance diminuera moins vite que $1/L$.

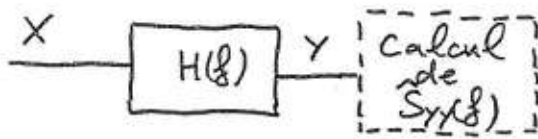
La contre partie de la réduction de la variance est une perte de résolution et un biais. En effet,

$$\begin{aligned}E[\hat{S}_{xx}^{MOY}(f)] &= \frac{1}{L} \sum_{i=0}^{L-1} E[\hat{S}_{xx}^{PER(i)}(f)] \\ &= E[\hat{S}_{xx}^{PER(i)}(f)] \\ &= T_K(f) * S_{xx}(f).\end{aligned}$$

La largeur du lobe principal est en $\frac{1}{K}$ alors qu'elle est en $\frac{1}{N}$ si on utilise un périodogramme sur toutes

Les données. La résolution fréquentielle est ainsi détériorée d'un facteur $L = N/K$.

Notons au passage une technique utile pour réduire le biais: il s'agit de « pré-blanchir » le signal afin de limiter les variations rapides de la dsp. Le dispositif est le suivant:



$$S_{yy}(f) = |H(f)|^2 S_{xx}(f),$$

$$\text{et } \hat{S}_{xx}(f) = \frac{\hat{S}_{yy}(f)}{|H(f)|^2}.$$

On utilise donc un filtre pour modéliser (en partie) la dsp. Cette idée sera reprise dans le cadre de la modélisation spectrale, où l'on cherchera effectivement à blanchir le signal: la dsp en sortie sera alors simplement $S_{yy}(f) = \sigma^2$, et

$$\hat{S}_{xx}(f) = \frac{\sigma^2}{|H(f)|^2}.$$

La dsp s'exprimera donc en fonction des paramètres du modèle du filtre.

Périodogramme lissé

Nous avons introduit le périodogramme lissé lors de l'étude du corrélogramme. Rappelons simplement que le lissage du périodogramme consiste à introduire une fenêtre spectrale $W(f)$ et à former

$$\hat{S}_{xx}^{\text{LIS}}(f) = W(f) * \hat{S}_{xx}^{\text{PER}}(f),$$

où il convient de choisir une fenêtre $W(f) \geq 0$. Le périodogramme lissé peut-être interprété comme la TF de la séquence de corrélation apodisée par $w(k)$.

En général, on choisit une fenêtre $w(k)$ paire, afin que sa TF $W(f)$ soit réelle, et lorsque le signal étudié est réel, on prend évidemment $w(k)$ réel. Il s'ensuit que $W(f)$ est elle-même réelle et paire.

Or si l'on applique directement le lissage en fréquence (i.e. sans passer par le biais du corrélogramme),

rien n'oblige à se restreindre à des fenêtres $w(f)$ paires. De plus, la fenêtre $w(k)$ doit être de longueur inférieure à $2N-1$ ($-(N-1)$ à $(N-1)$). Il est ainsi impossible de choisir une fenêtre en sinus cardinal alors que l'on peut, très bien employer une fenêtre rectangulaire pour le lissage. Tous les périodogrammes lissés ne peuvent donc pas être interprétés comme un corrélogramme.

La méthode de Welch

La méthode de Welch est une variante de la méthode du périodogramme moyenné dans laquelle :

- on utilise une fenêtre de pondération $w(k)$ sur chacun des segments (la méthode originale utilise une fenêtre en cosinus ajusté surelevé)
- on multiplie le nombre de segments en introduisant un recouvrement entre les segments

Avec un recouvrement de 50%, la variance est de $\frac{9}{16} \times \frac{1}{L} \text{Var} \{ S_x^{\text{PER}} \}$ ou L est le nombre de segments de même longueur pour le périodogramme moyenné. Au delà de 50% de recouvrement, on ne gagne plus en variance.

Exercice B -

On considère un filtre de réponse impulsionnelle $h(n) = a^n$, pour $n \geq 0$, attaqué par un bruit blanc $x(n)$ de variance σ_x^2 , et de sortie $y(n)$.

- a- Calculez l'autocorrélation théorique $R_{yy}(k)$. [vous utiliserez la formule des interférences].
- b- Dédisez en, par TF, la dsp. Tracez la.
- c- Comparez ce résultat au résultat obtenu à l'aide de la formule des interférences dans le domaine fréquentiel.
- d- Identifiez le coefficient d'un modèle AR(1) pour le signal $y(n)$. Donnez la variance du bruit génératrice
- e- Trouvez les coefficients d'un AR(e).

Exercice 4 -

On définit $\hat{S}_{xx}^{Luis}(f) = W(f) * \hat{S}_{xx}^{(B)}(f)$.
 En utilisant le fait que $\hat{S}_{xx}^{(B)}(f)$ est approximativement (et asymptotiquement) non biaisé, montrez que $E[\hat{S}_{xx}^{Luis}(f)] \sim W(f) * S_{xx}(f)$.

Envisagez le cas où x est un bruit blanc de variance σ_x^2 . Sous quelle condition sur $W(f)$ l'estimateur est-il non biaisé? Comment ceci se traduit-il sur $w(k) = TF^{-1}\{W\}$?

Exercice 1 - Pavé Temps-Fréquence

On note $(\Delta t)^2 = \frac{1}{E_x} \int (t-\bar{t})^2 |x(t)|^2 dt$ $\bar{t} = \int t |x(t)|^2 dt / E_x$
 $(\Delta f)^2 = \frac{1}{E_x} \int (f-\bar{f})^2 |X(f)|^2 df$ $\bar{f} = \int f |X(f)|^2 df / E_x$

les "épanouissements" temporel et fréquentiel, avec $E_x = \int |x(t)|^2 dt = \int |X(f)|^2 df$.

On peut montrer que $\Delta t \Delta f \geq \frac{1}{4\pi}$ (cf cours)

a) Montrez que $\Delta t \Delta f$ est invariant par « dilatation » temporelle ou fréquentielle. On pourra prendre $\bar{t} = \bar{f} = 0$

Indication: $TF\{x(at)\} = \frac{1}{|a|} X\left(\frac{f}{a}\right)$

b) Montrez que Δt et Δf sont invariants par décalage temporel ou fréquentiel. Considérez par exemple $y(t) = x(t-\tau)e^{j2\pi f_0 t}$.

Exercice 2 - Identification à partir du bispectre - (cas réel)

a) Proposer un algorithme permettant d'identifier la fonction de transfert $H(f)$ à partir du bispectre de la sortie d'un filtre excité par un bruit blanc

$B(f_1, f_2) = \int_3 H(f_1) H(f_2) H^*(f_1 + f_2)$

Indication: Travailler avec $B(f_1, -f_2)$ dans le cas réel $\rightarrow H(-f) = H^*(f)$

b) Montrez que $H(f)$ ne peut être identifié qu'à un facteur et une phase linéaire...

Exercice 5

On pondère les données $x(n)$ par une fenêtre $w(n)$, de longueur inférieure à N , le nombre de données disponibles. Montrez que la fonction d'autocorrélation, estimée à l'aide de l'estimateur biaisé, vaut en moyenne :

$$E[\hat{R}_{xx}] = w(k) \cdot R_{xx}(k) \quad \text{sur la longueur de } w(k) \text{ et 0 ailleurs.}$$

$w(k)$ est une fenêtre, dont vous donnerez l'expression en fonction de $P(n)$.
Montrez que $w(k)$ est ≥ 0 .

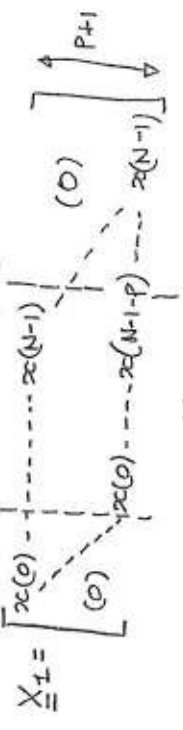
Exercice 6

On dispose de $N=8$ échantillons d'une sinusoïde de fréquence f_0 , échantillonnée à $F_e=4f_0$. La phase à l'origine est nulle, et l'amplitude vaut 1.

- a- Donnez les 8 échantillons $x(0) \dots x(7)$
- b- Calculez le périodogramme de x , et représentez-le.
- c- On introduit deux estimateurs de la matrice de corrélation, sous la forme

$$\hat{R}_1 = X_1 X_1^T \quad \text{et} \quad \hat{R}_2 = X_2 X_2^T$$

avec



X_2

Calculez \hat{R}_1 et \hat{R}_2 pour $P=2$, et calculez les paramètres d'un AR(2) dans les deux cas. Tracez la disp dans le second cas.

1D Traitement du signal - Essai corrigé
Remis par JF BERGER

Exercice 1 -

a) on considère $y(t) = x(at)$

$$(\Delta t)_{xy}^2 = \frac{1}{E_{xy}} \int_{-\infty}^{+\infty} t^2 |y(t)|^2 dt$$

$$E_{xy} = \int_{-\infty}^{+\infty} |y(t)|^2 dt = \int_{-\infty}^{+\infty} |x(at)|^2 dt$$

On pose $t' = at$
 $dt' = a dt$ } $E_{xy} = \int_{-\infty}^{+\infty} |x(t')|^2 \frac{dt'}{a} = \frac{E_x}{a}$

$$\int_{-\infty}^{+\infty} t^2 |x(at)|^2 dt = \int_{-\infty}^{+\infty} \frac{t'}{a^2} |x(t')|^2 \frac{dt'}{a} = \frac{1}{a^3} \int_{-\infty}^{+\infty} t'^2 |x(t')|^2 dt'$$

D'où $(\Delta t)_{xy}^2 = \frac{1}{a^2} (\Delta t)_x^2$

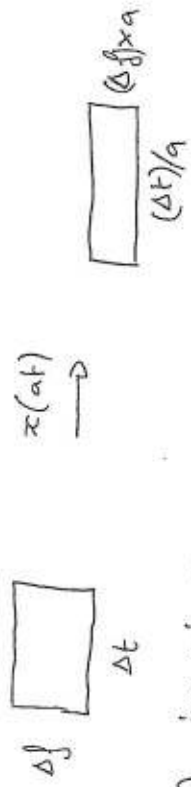
D'autre part, TF $\{x(at)\} = \frac{1}{|a|} X\left(\frac{f}{a}\right)$

$$(\Delta f)_{xy}^2 = \frac{1}{E_{xy}} \int_{-\infty}^{+\infty} f^2 \frac{1}{a^2} |X(f/a)|^2 df$$

$$= a^2 (\Delta f)_x^2$$

et finalement

$$(\Delta t)_{xy} (\Delta f)_{xy} = (\Delta t)_x (\Delta f)_x$$



b) invariance

$$y(t) = x(t-\tau) e^{j2\pi f_0 t}$$

$$(\Delta t)_{xy}^2 = \int_{-\infty}^{+\infty} (t-\tau)^2 |x(t-\tau)|^2 dt =$$

$$E_{xy} = \int_{-\infty}^{+\infty} t |x(t-\tau)|^2 dt = \int_{-\infty}^{+\infty} (t'+\tau) |x(t')|^2 dt' = \bar{t}_x + \tau$$

On obtient donc

$$(\Delta t)_{xy}^2 = \int_{-\infty}^{+\infty} (t-\tau)^2 |x(t-\tau)|^2 dt = \int_{-\infty}^{+\infty} (t'-\tau)^2 |x(t')|^2 dt' = (\Delta t)_x^2$$

De la même manière,

$$(\Delta f)_{xy}^2 = \int_{-\infty}^{+\infty} (f-f_0)^2 |X(f-f_0)|^2 df = (\Delta f)_x^2$$

Ainsi les élargissements temporels et fréquentiels ne dépendent pas de la position du pavé dans le plan temps-fréquence.

Exercice 2

On suppose $\gamma_3 = 1$ (coefficient absorbé dans H).

On a $H(f_1) = \frac{B(f_1, -f_2)}{H^*(f_2) H^*(f_1 - f_2)} = H^*(-f_1)$

Si $H(f)$ connu pour $f < f_2$ et $f_1 > f_2$, l'équation précédente permet de déterminer $H(f_1)$.

On dispose même de plusieurs équations redondantes [le équations] et l'on peut appliquer une procédure itérative. $H(-f_1)$ est obtenu par symétrie hermitienne condition initiale -

- $B(0,0) = |H(0)|^2 H(0)$ avec $H(0)$ réel

donc $H(0) = \text{sign } B(0,0) \times \sqrt[3]{|B(0,0)|}$

⇒ La phase de $H(1)$ ne peut pas être déterminée. On prend donc une phase arbitraire, qui entraînerait ensuite une ambiguïté...

1D Traitement du signal - Essai corrigé
Remis par JF BÉCHER

Exercice 1 -

a) on considère $y(t) = x(at)$

$$(\Delta t)_y = \frac{1}{E_y} \int_{-\infty}^{+\infty} t^2 |y(t)|^2 dt$$

$$E_y = \int_{-\infty}^{+\infty} |y(t)|^2 dt = \int_{-\infty}^{+\infty} |x(at)|^2 dt$$

on pose $t' = at$
 $\left. \begin{aligned} dt' = a dt \\ E_y = \int_{-\infty}^{+\infty} |x(t')|^2 \frac{dt'}{a} = \frac{E_x}{a} \end{aligned} \right\}$

$$\int t^2 |x(at)|^2 dt = \int \frac{t'}{a^2} |x(t')|^2 \frac{dt'}{a} = \frac{1}{a^3} \int t'^2 |x(t')|^2 dt'$$

D'où $(\Delta t)_y^2 = \frac{1}{a^2} (\Delta t)_x^2$

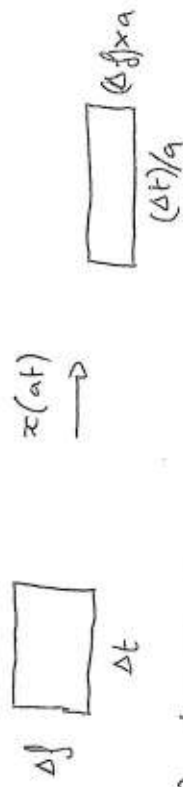
D'autre part, TF $\{x(at)\} = \frac{1}{|a|} X\left(\frac{f}{a}\right)$

$$(\Delta f)_y^2 = \frac{1}{E_y} \int f^2 \frac{1}{a^2} |X(f/a)|^2 df$$

$$= a^2 (\Delta f)_x^2 \quad \text{par les mêmes calculs}$$

et finalement

$$(\Delta t)_y (\Delta f)_y = (\Delta t)_x (\Delta f)_x$$



b) invariance

$$y(t) = x(t-\tau) e^{j2\pi f_0 t}$$

$$(\Delta t)_y^2 = \int (t-\tau)^2 |x(t-\tau)|^2 dt = \int (t'+\tau)^2 |x(t')|^2 dt' = \bar{E}_x + \tau^2$$

$$E_y = \int |x(t-\tau)|^2 dt = \int (t'+\tau)^2 |x(t')|^2 dt' = \bar{E}_x + \tau^2$$

On obtient donc

$$(\Delta t)_y^2 = \int (t-\tau)^2 |x(t-\tau)|^2 dt = \int (t'-\tau)^2 |x(t')|^2 dt' = (\Delta t)_x^2$$

De la même manière,

$$(\Delta f)_y^2 = \int (f-f_0)^2 |X(f-f_0)|^2 df = (\Delta f)_x^2$$

Ainsi, les élargissements temporels et fréquentiels ne dépendent pas de la position du pavé dans le plan temps-fréquence.

Exercice 2 -

On suppose $\chi_3 = \pm 1$ (coefficient absorbé dans H).

On a $H(f_1) = \frac{H^*(f_2) H^*(f_1 - f_2)}{B(f_1, -f_2)} = H^*(-f_1)$

Si $H(f)$ connu pour $f < f_2$ et $f_1 > f_2$, l'équation précédente permet de déterminer $H(f_1)$.
 On dispose même de plusieurs équations redondantes [2 équations] et l'on peut appliquer une procédure itérative. $H(-f_1)$ est obtenu par symétrie hermitienne

Condition initiale -

$$B(f_1, 0) = |H(f_1)|^2 H(f_1) \quad \text{avec } H(0) \text{ réel}$$

$$\text{donc } H(0) = \text{sign } B(f_1, 0) \times \sqrt[3]{|B(f_1, 0)|}$$

La phase de $H(f_1)$ ne peut pas être déterminée. On prend donc une phase arbitraire, qui entraînerait ensuite une ambiguïté en phase linéaire sur Ω .

b) ambiguïté.

Le coefficient d'asymétrie, χ_3 , est inconnu. On ne peut que l'inclure dans l'identification de $H(f)$, qui du coup est à un facteur près.

Si on pose $\tilde{H}(f) = H(f)e^{-j2\pi f t_0}$, le bispectre devient

$$\begin{aligned} \tilde{B}(f_1, f_2) &= \chi_3 \tilde{H}(f_1) \tilde{H}(f_2) \tilde{H}^*(f_1 + f_2) \\ &= \chi_3 H(f_1) H(f_2) H^*(f_1 + f_2) e^{-j2\pi(f_1 + f_2)t_0} \\ &= B(f_1, f_2) \quad \left[e^{+j2\pi(f_1 + f_2)t_0} \right] \end{aligned}$$

et est donc inchangé.

On ne pourra identifier $H(f)$ qu'à une phase linéaire près, c'est-à-dire $h(t)$ qu'à un retard près.

Exercice 3.-

a- Formule des interférences (Temps):

$$R_{YY}(k) = (h * R_{XX} * h^*)(k)$$

Soit $R_{XX}(k) = \sigma^2$, $h(n) = a^n$, $h^*(n) = a^{-n}$ ($n < 0$)
 $(h * h^*)(k) = \sum_{n=0}^{+\infty} a^n a^{n-k}$ pour $k < 0$

$$= a^{-k} \sum_{n=0}^{+\infty} a^{2n} = a^{-k} \times \frac{1}{1-a^2}$$

On en déduit que $R_{YY}(k) = a^{|k|} \frac{\sigma^2}{1-a^2}$.

b- Calcul de la dsp:

$$S_{YY}(f) = \sum_{k=-\infty}^{+\infty} R_{YY}(k) e^{-j2\pi f k}$$

$$= \frac{\sigma^2}{1-a^2} \left[\sum_{k=0}^{+\infty} a^k e^{-j2\pi f k} + \sum_{k=0}^{+\infty} a^k e^{j2\pi f k} \right]$$

$$= \frac{\sigma^2}{1-a^2} \left[\frac{1}{1-a e^{-j2\pi f}} + \frac{1}{1-a e^{j2\pi f}} - 1 \right]$$

$$= \frac{\sigma^2}{|1-a e^{-j2\pi f}|^2} = \frac{\sigma^2}{(1-a^2) - 2a \cos(2\pi f)}$$

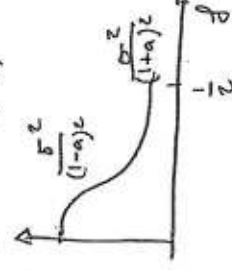
c- Formule des interférences (fréquence)

$$S_{YY}(f) = \sigma^2 |H(f)|^2$$

or $H(f) = \sum_{n=0}^{+\infty} a^n \delta^{-n} = \frac{1}{1-a \delta^{-1}}$

D'où $H(f) = \frac{1}{1-a e^{-j2\pi f}}$

et $S_{YY}(f) = \frac{\sigma^2}{|1-a e^{-j2\pi f}|^2}$



d. Coefficients AR(1) - Equations normales

$$\begin{cases} R_{YY}(k) = a R_{YY}(k-1) & k \neq 0 \\ R_{YY}(0) = a R_{YY}(-1) + \epsilon^2 & (k=0) \end{cases}$$

d'où $a = \frac{R_{YY}(k)}{R_{YY}(k-1)}$ $\epsilon^2 = \frac{\sigma^2(1-a^2)}{(1-a^2)} = \sigma^2$

e. Coefficients AR(1)
 $R(0) = \sigma^2/(1-a^2)$ $R(1) = \sigma^2 a/(1-a^2)$ $R(2) = \frac{a^2 \sigma^2}{1-a^2}$
 Equations normales:

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

$$\Rightarrow \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \frac{1}{1-a^2} \begin{bmatrix} 1 & -a \\ -a & 1 \end{bmatrix} \begin{bmatrix} a \\ a^2 \end{bmatrix} = \begin{bmatrix} a \\ 0 \end{bmatrix}$$

Logique!
 Le processus est AR(1).

Exercice 4

$$\hat{S}_{XX}^{Lis}(f) = \int_{[1]}^{[2]} W(u) \hat{S}_{XX}(f-u) du$$

$$E[\hat{S}_{XX}^{Lis}(f)] = \int_{[1]}^{[2]} W(u) E[\hat{S}_{XX}(f-u)] du$$

$$= (W * S_{XX})(f)$$

Si X est un bruit blanc, $E[\hat{S}_{XX}(f)] = \sigma^2$, et
 $E[\hat{S}_{XX}^{Lis}(f)] = \sigma^2 \int_{[1]}^{[2]} W(u) du$

Pour éviter de biaiser l'estimateur, il faut donc que
 $\int_{[1]}^{[2]} W(u) du = 1$

En notant que $\int W(f) df = \int W(f) e^{j2\pi f \cdot 0} df$
 $= TF^{-1}\{W(f)\}|_{k=0} = W(0)$,
 on trouve qu'il faut que $w(0) = 1$

Exercice 5:

$$x(n) \rightarrow x_p(n) = p(n) \cdot z(n)$$

$$\hat{R}_{XX}(k) = \frac{1}{N} \sum_{n=0}^{N-1} p(n) p(n-k) z(n) z(n-k)$$

soit $E[\hat{R}_{XX}(k)] = \underbrace{\frac{1}{N} \sum_{n=0}^{N-1} p(n) p(n-k)}_{w(k)} R_{ZZ}(k) = w(k) \cdot R_{ZZ}(k)$

On observe donc que $w(k) = (p * p^*)(k)$
 Puisque TF $\{p(n)\} = P(f)$, on a $w(f) = |P(f)|^2 \geq 0 \dots$

Exercice 6:

$$x(n) = \sin\left(\frac{2\pi f_0 n}{4f_s}\right) = \sin\left(\frac{\pi}{2} n\right)$$

$$x(0) = 0 \quad x(1) = 1 \quad x(2) = 0 \quad x(3) = -1 \quad x(4) = 0$$

$$x(5) = 1 \quad x(6) = 0 \quad x(7) = -1$$

La transformée de Fourier de x vaut:

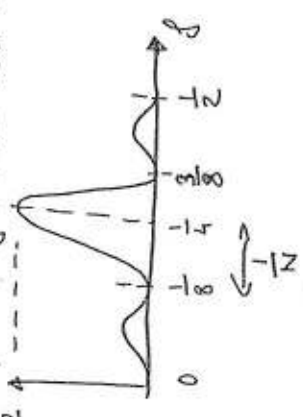
$$X(f) = \sum_{n=0}^7 x(n) e^{-j2\pi f n} = e^{-j2\pi f} + e^{-j6\pi f} - e^{-j10\pi f} - e^{-j14\pi f}$$

$$X(f) = e^{-j2\pi f} \sum_{k=0}^3 (-1)^k e^{-j4\pi k f}$$

$$= e^{-j2\pi f} \cdot \frac{1 - e^{-j16\pi f}}{1 + e^{-j4\pi f}}$$

$$= e^{-j8\pi f} \frac{\sin(8\pi f)}{\cos(2\pi f)}$$

le periodogramme vaut donc $S_{xx}(f) = \frac{1}{8} \left(\frac{\sin 8\pi f}{\cos 2\pi f} \right)^2$



$$\hat{R}_1 = \begin{bmatrix} 4 & 0 & -3 \\ 0 & 4 & 0 \\ -3 & 0 & 4 \end{bmatrix} \quad \hat{R}_2 = \begin{bmatrix} 3 & 0 & -3 \\ 0 & 3 & 0 \\ -3 & 0 & 3 \end{bmatrix}$$

Le calcul des paramètres AR(z) consiste à résoudre

$$\hat{R}_2 \begin{bmatrix} 1 \\ -a_1 \\ -a_2 \end{bmatrix} = \begin{bmatrix} \epsilon^2 \\ 0 \\ 0 \end{bmatrix}$$

On obtient:

pour \hat{R}_1 : $a_1 = 0$
 $a_2 = -3/4$
 $\epsilon^2 = 7/4$

pour \hat{R}_2 : $a_1 = 0$
 $a_2 = -1$
 $\epsilon^2 = 0$

On observe que $|1 - a_1 e^{-j2\pi f} - a_2 e^{-j4\pi f}|^2$
 $= |1 + e^{-j4\pi f}|^2$

s'annule pour $f = 1/4 \Rightarrow S_{xx}(f=1/4) = +\infty$