

Introduction aux signaux discrets aléatoires

TP N° II

I-Notions théoriques intervenant dans le TP

I.1 Variables aléatoires

Variables aléatoire (VA) réelles (notions supposées connues)

Pour une VA : loi, moyenne et variance.

Pour deux VA : loi bi-dimensionnelle, indépendance, coefficient de corrélation.

I.2 Suites aléatoires réelles (signaux aléatoires discrets SAD).

Généralités

Définition : $X[n], n \in \mathbb{Z}$ est une suite aléatoire si toutes les $X[n]$ sont des VA (sur un même espace probabilisé). Suivant le contexte, ces variables aléatoires peuvent être considérées comme à valeurs dans $\mathbf{R}$ (corps des réels) ou dans $\mathbf{C}$ (corps des complexes). Dans ce TP les VA représentant les amplitudes des signaux temporels (temps discret) seront à valeurs réelles. Tout symbole de conjugaison complexe portant sur ces VA les laissera donc inchangées.

Etude à l'ordre 1 : comportement statistique des $X[n]$ prises individuellement (loi, moyenne et variance).

Etude à l'ordre 2 : comportement statistique des couples $(X[n], X[m])$ où n et m sont arbitraires (loi, coefficient de corrélation). Les suites à deux indices entiers dites respectivement fonction de covariance discrète non centrée et fonction de covariance discrète centrée sont en particulier introduites :

$$C_X[n, m] = E(X[n]X^*[m]), \quad C_{X_C}[n, m] = E(X_C[n]X_C^*[m]), \quad (n, m) \in \mathbb{Z}^2$$

Suites aléatoires stationnaires au sens large

Définition : la suite $X[n], n \in \mathbb{Z}$, est stationnaire au Sens Large (SSL) si :

$$\forall n E(X^2[n]) < \infty$$

$$\forall n E(X[n]) = m_X = \text{constante}$$

$$\forall n \forall p E(X[n]X[n-p]) = C_X[n, m] = \Gamma_X[p] = \text{fonction de } p \text{ seul}$$

La suite $(\Gamma_X[p])$ est appelée *fonction d'autocorrélation discrète*. On démontre qu'elle est paire si les $X[n]$ sont réelles et que $|\Gamma_X[p]| \leq \Gamma_X(0)$ (ce qui permet de contrôler des résultats).

La suite des coefficients de corrélation : pour deux instants n et $n-p$ et pour une suite SSL, le coefficient de corrélation entre les VA $X[n]$ et $X[n-p]$ ne dépend que de p et est égal à :

$$\rho_X(p) = \frac{\Gamma_X[p] - m_X^2}{\Gamma_X[0] - m_X^2}$$

Considérations énergétiques :

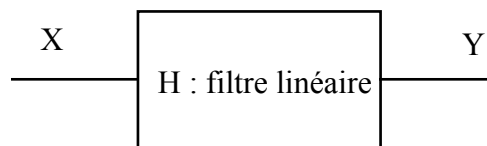
$E[X(n)^2]$ s'interprète comme la puissance statistique moyenne (PSM(X)) de X à l'instant n . Cette quantité ne dépend pas de n si X est SSL. Elle est alors égale à $\Gamma_X[0]$. On introduit la *densité spectrale de puissance (DSP)*, notée γ_X^d , pour X SSL, comme étant la TFS (Transformée de Fourier au sens des suites) de Γ_X . On a donc :

$$\gamma_X^d = \sum_{p=-\infty}^{+\infty} \Gamma_X[p] e^{-2\pi j p f}, \quad \Gamma_X[p] = \int_{-0.5}^{0.5} \gamma_X^d(f) e^{2\pi j p f} df, \quad p \in \mathbb{Z}, \quad \Gamma_X[0] = \int_{-0.5}^{0.5} \gamma_X^d(f) df$$

et si X s'exprime en volts alors $\Gamma_X(0)$ s'exprime en volts au carré et γ_X^d s'exprime en volts au carré par hertz. On montre que γ_X^d (qui est bien sûr périodique de période 1) est toujours réelle, positive et paire pour une suite réelle. Intuitivement, à une suite X d'allure "régulière" sans variations brusques, correspondra une fonction γ_X^d à "support concentré" autour de l'origine. C'est le théorème de Wiener-Kintchine présenté en cours qui justifie de prendre comme définition de la DSP la TFS de la fonction de corrélation discrète.

Bruit blanc discret : on appelle bruit blanc discret tout SAD SSL correspondant à une suite de VA décorréliées deux à deux : $\forall p \neq 0 \rho_X[p] = 0$. La fonction de corrélation d'un bruit blanc est donc complètement spécifiée par une moyenne commune à tous les échantillons et par une variance également commune à tous les échantillons.

Filtrage linéaire discret d'une suite aléatoire.



$X[n], n \in \mathbb{Z}$ étant une suite aléatoire, $Y[n] = (H * X)[n], n \in \mathbb{Z}$ (convolution entre X et H) est aussi une suite aléatoire. On montre alors que :

Tableau I

Si X est SSL:

- Y est SSL,
- $m_Y = m_X \hat{H}(0)$; $\hat{H} = TFS(H)$
- $\Gamma_Y[p] = (COR(H, H) * \Gamma_X)[p]$
- $\gamma_Y^d(f) = |\hat{H}(f)|^2 \gamma_X^d(f)$

I.3 Comment mesurer les caractéristiques statistiques d'un SAD à partir d'une réalisation ?

Problème posé

On dispose d'une réalisation $X_\omega[0], X_\omega[1], \dots, X_\omega[N-1]$ de la restriction du signal aléatoire discret $X[n]$ aux instants $n=0 \dots n=N-1$, ω marquant l'influence du "hasard" et N étant généralement grand. Les grandeurs statistiques que l'on désire mesurer sont supposées ici invariantes relativement à un décalage temporel arbitraire (autrement dit ces quantités sont supposées stationnaires).

Certaines quantités Θ_X statistiquement invariantes relativement à un décalage temporel (comme par exemple $\Theta_X = m_X$ ou encore $\Theta_X = \Gamma_X[p]$ pour une certaine valeur de p) et qui ont une signification théorique *a priori* (moyenne, densité de probabilité, fonction d'autocorrélation, ...) ne sont cependant pas connues numériquement par l'expérimentateur qui désire alors les mesurer à partir d'une réalisation à sa disposition.

Cette mesure correspondra à une réalisation $\tilde{\Theta}_X(\omega)$ d'une VA $\tilde{\Theta}_X$ construite sous la forme d'une fonction $h(X[0], \dots, X[N-1])$ des échantillons observés.

Sera ainsi introduite une erreur de mesure $\Theta_X - \tilde{\Theta}_X$ qui s'interprétera elle aussi comme une variable aléatoire. Une caractérisation statistique courante de cette erreur est :

$$\begin{aligned}\varepsilon_{\Theta_X}^2 &= E\left[\left(\Theta_X - \tilde{\Theta}_X\right)^2\right] \\ &= \left(E\left[\left(\Theta_X - \tilde{\Theta}_X\right)\right]\right)^2 + \text{VAR}(\tilde{\Theta}_X) \\ &= B_{\Theta_X}^2 + \text{VAR}(\tilde{\Theta}_X)\end{aligned}$$

B_{Θ_X} est appelé le biais de la mesure,

$\varepsilon_{\Theta_X}^2$ est appelée erreur quadratique moyenne.

Moins le biais et l'erreur quadratique moyenne seront élevés et meilleure sera la mesure.

Moyennes expérimentales

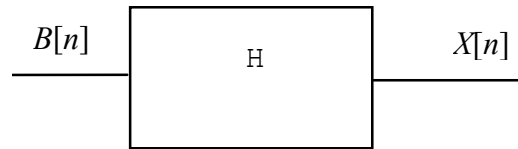
Une méthode générale pour obtenir des mesures est d'utiliser les moyennes expérimentales associées aux moyennes statistiques recherchées comme par exemple :

Tableau II

Grandeur mesurée	Moyenne statistique	Moyenne expérimentale
$\Theta_X = m_X$	$E[X[n]]$	$\tilde{\Theta}_X = \bar{X}_N = \frac{1}{N} \sum_{n=0}^{N-1} X[n]$
$\Theta_X = \text{VAR}(X)$	$E[X[n]^2] - m_X^2$	$\tilde{\Theta}_X = \bar{X}_{cN}^2 = \frac{1}{N} \sum_{n=0}^{N-1} X(n)^2 - \bar{X}_N^2$
$\Theta[p] = \Gamma_X[p]$	$E[X[n]X[n-p]]$	$\tilde{\Theta}_X = \overline{XX}_N[p] = \frac{1}{N-p} \sum_{n=p}^{N-1} X[n]X[n-p]$
$\Theta_X[p] = \rho_X[p]$	$\frac{\Gamma_X(p) - m_X^2}{\Gamma_X(0) - m_X^2}$	A votre avis ?

II. Manipulation

On considère le système suivant :



- B est une suite de VA décorréélées deux à deux de même loi gaussienne $\mathcal{N}(m_B, \sigma_B^2)$.
- $\mathcal{H}$ est une filtre linéaire et homogène de réponse impulsionnelle :

$$H : b \cdot a^n U[n] \xrightarrow{\text{tz}} H(z) = \frac{b}{1 - az^{-1}}, |z| > |a|, |a| < 1$$

Le but de ce TP est de :

- simuler sous matlab B et X ,
- appliquer les formules (moyennes temporelles sur une réalisation) données dans le tableau II pour mesurer les moyennes et fonctions d'autocorrélation (théoriques) pour B et X ,
- obtenir une mesure des DSP pour B et X ,
- comparer ces mesures aux quantités théoriques correspondantes (celles du tableau I).

Préparation

1- On se reportera à l'annexe 1 pour étudier comment simuler des VA de loi gaussienne.

2- Expliquer pourquoi l'opération de filtrage peut se réaliser avec l'algorithme :

$$X(n) = aX(n-1) + bB(n), n \text{ positif ou nul, avec } X[-1] = \text{condition initiale imposée.}$$

3- Appliquer la définition de la moyenne et de la fonction d'autocorrélation et utiliser les formules du tableau I pour calculer :

$$\Gamma_X[p] = ?, \gamma_X^d(f) = ? \quad \Gamma_B[p] = ?, \gamma_B^d(f) = ?$$

Donner l'allure graphique de ces résultats.

4- Le biais est-il non nul pour les mesures de la colonne de droite du tableau II ?

5- On choisit de prendre $a = 0,9$ et $b = 2$. Calculer m_B et σ_B^2 pour avoir $m_X = 1$ et $\text{VAR}(X[n]) = 2$.

6- Ecrire, en utilisant le tableau II, des procédures matlab qui, à partir d'un tableau contenant une réalisation de N échantillons de la suite $X[n]$ supposée SSL, fournissent une mesure (une estimée sous forme de moyenne expérimentale) $\bar{X}$ de sa moyenne et une mesure $\overline{XX}[p]$ de son autocorrélation pour p allant de zéro à $p_M, p_M < N$.

7- Ecrire une procédure qui, à partir d'un tableau $\overline{XX}[p]$ pour p allant de zéro à p_M , calcule une mesure $\tilde{\gamma}_X^d(k/N), k = 0 \dots N-1$ de la densité spectrale de puissance $\gamma_X^d(f)$, pour $f = k/N, k = 0 \dots N-1$ (cad pour N valeurs de fréquence réparties entre 0 et 1) définie par :

$\tilde{\gamma}_X^d(k/N) = \text{FFT}(T)(k), k \in 0, \dots, N-1$ et où la suite $T[p], p$ allant de zéro à $N-1$, est définie par :

$$\begin{cases} T[p] = 0 & \text{si } p_M + 1 \leq p \leq N-1 - p_M \\ T[p] = \overline{XX}[p]F[p] & \text{si } 0 \leq p \leq p_M \\ T[N-p] = \overline{XX}[p]F[p] & \text{si } 1 \leq p \leq p_M \end{cases}$$

$F[p], p$ allant de zéro à p_M , est une fenêtre de pondération définie par :

$$F[p] = 0.42 + 0.5 \cdot \cos\left(\frac{\pi p}{p_M}\right) + 0.08 \cdot \cos\left(\frac{2\pi p}{p_M}\right); 0 \leq p \leq p_M$$

$\tilde{\gamma}_X^d$ correspond à une tableau indexé de 0 à $N-1$ (I à N en matlab). N étant le nombre de points utilisé par la FFT devra donc être une puissance de 2.

Remarque : intuitivement l'introduction de $\tilde{\gamma}_X^d$ pour mesurer la DSP de X se justifie en considérant que cela revient à calculer la TFS d'une estimation de la fonction de corrélation discrète.

Mise en oeuvre en matlab

On utilisera comme base de travail les valeurs de a, b, m_B et $\text{VAR}(B)$ introduites plus haut. Ces valeurs pourront cependant être changées à loisir pour examiner des variantes ($m_B = 0, \dots$).

- 1- Simuler et visualiser la réponse du filtre $\mathcal{H}$, avec une condition initiale $X[-1] = 0$, aux entrées $E[n] = \delta_0[n]$ et $E[n] = U[n]$.
- 2- Générer plusieurs réalisations de B et de X et les visualiser pour $N = 512$.
- 3- Pour plusieurs couples de réalisations de (B, X) et pour différentes valeurs de N (128, 256, 1024, ...), calculer les réalisations des moyennes et autocorrélations expérimentales $\overline{X}, \overline{B}, \overline{XX}[p]$ et $\overline{BB}[p]$ de X et de B pour p allant de zéro à $p_M = 30$. **On observera la variabilité des résultats obtenus** en fonction de la réalisation, de N , et par rapport aux quantités théoriques correspondantes. Pour cela, on pourra évaluer, au moyen de moyennes expérimentales sur 20 réalisations, les valeurs de B_{Θ_X} et $\varepsilon_{\Theta_X}^2$ pour $\Theta_X = \Gamma_X[p]$ et $\Theta_X = \Gamma_X[p]$.
- 4- Représenter un histogramme des amplitudes pour les signaux d'entrée et de sortie. Interpréter la distribution des amplitudes en sortie du filtre.
- 5- Représenter à l'écran, pour $p = 0, 8, 15$, l'ensemble des points de coordonnées

($x = B[n], y = B[n-p]$) pour $n = p, p+1, \dots, N-1$. Recommencer avec X à la place de B et interpréter en considérant les valeurs de $\rho_X[p], p = 0, 8, 15$.

- 6- Proposer et appliquer un principe de mesure de la suite des coefficients de corrélation pour B et X .
- 7- Pour la fenêtre de pondération spécifiée plus haut, calculer les mesures des densités spectrales de B et X pour plusieurs réalisations. Comparer avec les valeurs théoriques γ_B^d et γ_X^d que l'on obtiendra en prenant les TFS de Γ_B et Γ_X .
- 8- Proposer et appliquer un principe de mesure de la suite des coefficients d'intercorrélation entre l'entrée et la sortie du filtre.

Annexe 1 : Génération de nombres aléatoires

On considère ici le principe de la génération d'une suite $X(n)$ de V.A indépendantes de même loi P_X .

On utilise la primitive *rand* de matlab qui retourne un nombre aléatoire de loi uniforme entre zéro et un. *Des tirages successifs sont considérés indépendants.*

Loi uniforme :

$$P_X(u) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq u \leq b \\ 0 & \text{sinon} \end{cases} ; \text{ on fait } X = a + (b-a) * \text{rand}.$$

Loi exponentielle de paramètre a :

$$P_X(u) = \begin{cases} a \cdot \exp(-au) & \text{si } 0 \leq u \\ 0 & \text{sinon} \end{cases} ; \text{ on fait } X = \frac{-\text{Log}(\text{rand})}{a}.$$

Loi normale centrée et de variance unité :

On veut obtenir $P_X(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$. Pour cela, on utilise deux tirages successifs (donc indépendants) de *rand* pour générer deux tirages indépendants de X :

- $\Theta = 2\pi \cdot \text{rand}, U = \text{rand} \Rightarrow U$ et Θ sont deux V.A indépendantes,
- $X_1 = \sqrt{-2\text{Log}(U)} \cdot \text{Cos}(\Theta)$
- $X_2 = \sqrt{-2\text{Log}(U)} \cdot \text{Sin}(\Theta)$

On montre alors que X_1 et X_2 sont deux V.A indépendantes de même loi normale centrées et de variance unité. Pour générer $2N$ V.A de ce type on peut donc se contenter de répéter ces opérations N fois.

Annexe 2 : Correspondances entre les noms de variables dans les programmes à écrire et les paramètres théoriques.

Pour des raisons d'homogénéité entre les binômes, il est préférable d'utiliser les identificateurs suivants :

<i>Théorique</i>	<i>Identificateur utilisé</i>
N	nbp
Bruit B	B
X	X
$\bar{B}$	MoyB
$\bar{X}$	MoyX
p_M	nbpac
$\overline{BB}$	BB
$\overline{XX}$	XX
B_{Θ_X}	Biais
$\varepsilon_{\Theta_X}^2$	Eqm
Γ_X	acX
Γ_B	acB
ρ_X	ccX
$\gamma^d_B, \tilde{\gamma}^d_B$	DspB, DspeB
$\gamma^d_X, \tilde{\gamma}^d_X$	DspX, DspeX
$\Gamma_{BX}, \overline{BX}$	IcBX, BX