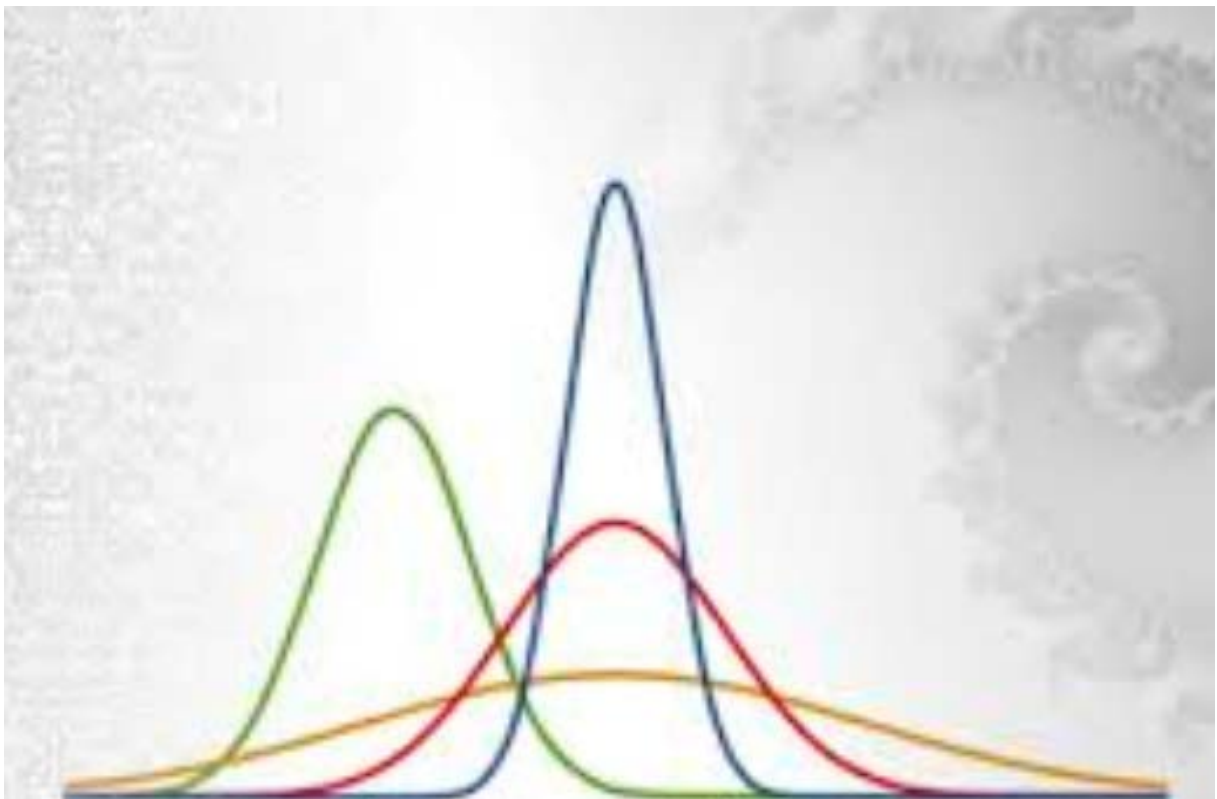

Variables aléatoires et processus stochastiques



Auteurs :

Sidi Mohammed Hadj IRID

Mourad HADJILA

Adardour Haroune Errachid

Laboratoire STIC

Sommaire

Introduction	1
TP 1 : Probabilité et nombres aléatoires	2
TP 2 : Processus aléatoires	9
TP 3 : Densité spectrale de puissance	23
TP 4 : Estimation de la DSP en utilisant les méthodes de Welch, Multitaper, et Burg	33

Introduction

Le concept de aléatoire est important dans presque tous les aspects des systèmes et des sciences de l'ingénierie modernes. Cela inclut l'informatique, la biologie, la médecine, les sciences sociales et les sciences de la gestion, entre autres. Chaque fois que nous sommes incapables de connaître le comportement futur exact de certains phénomènes, nous disons qu'il est aléatoire. En tant que tels, les phénomènes aléatoires ne peuvent être qu'approximés.

Dans de nombreuses circonstances pratiques, les phénomènes observés présentent des variations importantes, alors que les informations pertinentes n'ont pas changé. Ainsi, si vous enregistrez les signaux obtenus en prononçant plusieurs fois le son "A", une simple observation vous dira que tous les enregistrements sont différents même s'ils ont tous le même son. Il n'est ni facile ni pertinent d'essayer de trouver une équation déterministe pour décrire l'évolution dans le temps d'un tel phénomène. Le modèle utilisé pour décrire cette variabilité est basé sur le concept de variables aléatoires

Une variable aléatoire réelle peut être entièrement caractérisée par différents moyens: fonction de distribution cumulative (cdf), fonction de densité de probabilité (pdf), fonction caractéristique, moments, etc.

Un processus aléatoire est un ensemble de variables aléatoires $X(t)$ indexées dans le temps définies dans le même espace de probabilité. Si les valeurs possibles pour t appartiennent à \mathbb{R} , le processus est appelé processus aléatoire à temps continu. Si les valeurs possibles pour t appartiennent à \mathbb{Z} , nous avons affaire à un processus aléatoire à temps discret.

Les deux propriétés les plus importantes d'un processus aléatoire sont la stationnarité et l'ergodicité. La stationnarité de sens strict est liée à l'invariance dans le temps des statistiques de processus aléatoires.

Le présent polycopié est un recueil de travaux pratiques faits avec le logiciel MATLAB dédiés aux concepts de variables aléatoires et processus stochastiques. Ce polycopié est constitué de quatre TPs, dont l'organisation est comme suit. A travers le TP 1, nous expliquons ce qu'est une variable aléatoire et comment elle est décrite par la fonction de densité de probabilité. Nous avons besoin de certains paramètres pour la caractérisation de la variable aléatoire. Les paramètres les plus importants sont la valeur moyenne et la variance.

Le concept de processus aléatoire est introduit et expliqué dans le TP 2 via des exercices. Une description du processus aléatoire est donnée et ses caractéristiques importantes, telles que la stationnarité et l'ergodicité, sont expliquées dans la partie théorique. Une attention particulière est également accordée à une fonction d'autocorrélation ; son importance et ses propriétés sont expliquées.

Le travail pratique 3 a été consacré à l'estimation de la densité spectrale de puissance d'un signal en utilisant un périodogramme simple.

Le travail pratique 4 porte sur l'estimation de la densité spectrale de puissance d'un signal en utilisant les méthodes de Welch, Multitaper et Burg puis les comparer avec l'estimateur simple.

TP 1 Probabilité et nombres aléatoires

But : Ce TP a pour objectif de se familiariser avec les différentes fonctions MATLAB permettant de générer les valeurs aléatoires suivant les distributions de probabilité ainsi que le calcul des fonctions de densité de probabilité (PDF) et celles de distribution cumulée normale et inverse.

I. Valeurs aléatoires de base, permutations, et des échantillons

- 1) **rand** → Génère une valeur aléatoire uniforme continue entre 0 et 1.
rand(n, 1) → Générer un vecteur colonne de n valeurs aléatoires uniformes entre 0 et 1.
a+(b-a)*rand (1, n) → Génère un vecteur ligne de n valeurs aléatoires uniformes entre a et b.
- 2) **rand(m,n)** → Génère une matrice m × n de valeurs aléatoires uniformes entre 0 et 1.
- 3) **randn** → Génère des tableaux de nombres aléatoires dont les éléments sont normalement distribués avec une moyenne 0 et une variance 1.
- 4) **m+v*randn (1, n)** → Génère des tableaux de nombres aléatoires dont les éléments sont normalement distribués avec une moyenne **m** et une variance **v**.

Exemple 1 :

Générer un vecteur colonne r de 10 lignes de nombres aléatoires distribués uniformément dans l'intervalle [-5,5].

$$r = -5 + (5+5)*rand(10,1)$$

Exemple 2 :

Créer un vecteur de 1000 valeurs aléatoires tracées à partir d'une distribution normale avec une moyenne de 500 et un écart type (déviation standard) de 5.

$$a = 5;$$

$$b = 500;$$

$$y = a.*randn(1000,1) + b;$$

Rappel :

On rappelle que :

$$m=(a+b)/2$$

$$v=(b-a)^2/12$$

m représente la moyenne, v représente la variance, et a et b représentent respectivement les bornes inférieure et supérieure de l'intervalle [a b].

La variance est égale au carré de l'écart type :

$$v = \sigma^2$$

II. PDF, CDF et CDF inverse

PDFs est un sigle de Probability Density Functions, alors que CDFs est un sigle de Cumulative Distribution Functions. Dans MATLAB, ces fonctions font partie de la boîte à outils statistics.

Il existe une fonction PDF correspondant à chacune des fonctions qui génère des valeurs aléatoires. Elle peut être obtenue en changeant le "rnd" (pour "random") à la fin du nom de la fonction par «pdf». Le premier argument des fonctions PDF est la valeur à laquelle vous voulez évaluer la fonction ; les arguments suivants sont les paramètres de la distribution de probabilité. Par exemple, pour calculer la densité à $x = 0.2$ pour une distribution normale avec une moyenne de 1.3 et un écart type de 1.1, utilisez **normpdf (0.2, 1.3, 1.1)**.

De même, il existe une fonction CDF correspondant à chaque fonction de générateur de nombre aléatoire. Elle peut être obtenue en changeant le suffixe "rnd" dans le nom de la fonction par "cdf." Le premier argument des fonctions CDF est la valeur x à laquelle vous voulez évaluer la probabilité cumulée $P(X \leq x)$. Par exemple, pour calculer la probabilité cumulée $P(X \leq 1.3)$ pour une valeur aléatoire X qui suit une distribution exponentielle avec une moyenne de 0.7, utilisez **expcdf (1.3, 0.7)**.

Enfin, les fonctions CDF inverses peuvent être obtenues en modifiant le suffixe dans le nom de la fonction par "inv". Le premier argument des fonctions CDF inverses est la probabilité p ; la fonction trouve alors la valeur x satisfaisant $P(X \leq x) = p$. Par exemple, **expinv (0.7, 3)**.

III. Comment estimer la densité et la distribution dans MATLAB

III.1. Fonction de densité

Ici, nous considérons la variable uniforme dans l'intervalle [0, 1]. La PDF et la distribution sont présentées sur la Fig. 1.

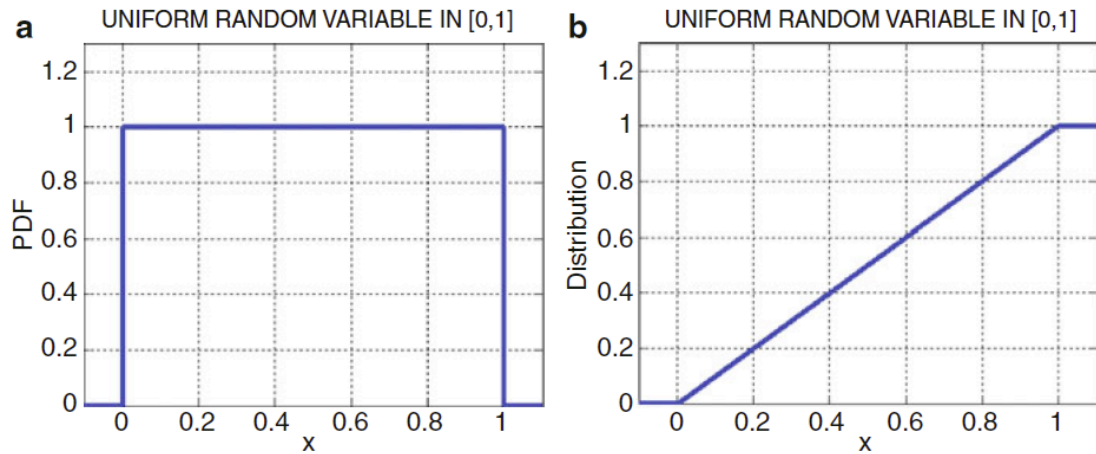


Fig.1 (a) PDF et (b) la distribution de la variable aléatoire uniforme dans l'intervalle [0, 1]

La variable uniforme dans l'intervalle [0, 1] peut être générée en utilisant le fichier Matlab rand.m. La figure 2.a montre la variable uniforme générée X avec $N=10000$ valeurs, tandis que la figure 2.b montre que les 1000 premiers échantillons. De la figure 2, il est évident pourquoi la variable est appelée «uniforme» (ses valeurs occupent uniformément toute la gamme).

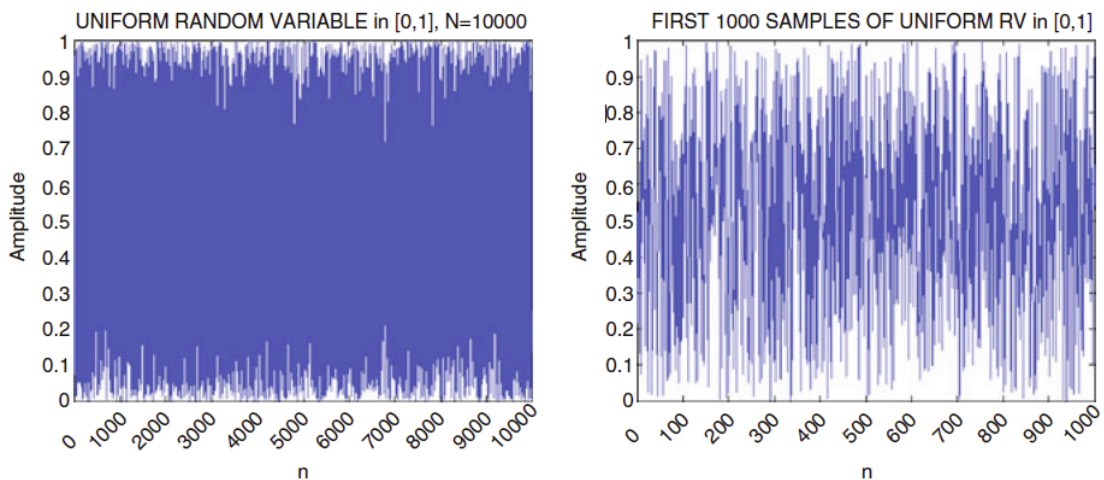


Fig.2 Variable uniforme générée en MATLAB

Afin d'estimer la PDF de la variable aléatoire X, on divise la plage de la variable [0, 1] en M cellules à distance égale Δx .

L'histogramme $NN = \text{hist}(x, M)$ indique les valeurs de N_i , $i = 1, \dots, M$ (i.e., le nombre de valeurs de la variable aléatoire X dans chaque cellule), où M est le nombre de cellules. Si l'on utilise $M = 10$, on obtient par exemple le résultat suivant :

$$NN = [952, 994, 993, 1012, 985, 1034, 1006, 998, 1047, 979] \quad (1)$$

Le résultat (1) montre que parmi les $N = 10\,000$ valeurs de la variable aléatoire, 952 se trouvent dans la première cellule, 994 dans la seconde cellule, et ainsi de suite. Nous pouvons également noter que le nombre de valeurs dans les cellules ne varie pas de manière significative.

Le tracé d'un histogramme peut être obtenu en tapant `hist(X, 10)` dans le prompt MATLAB. La figure 3 montre cet histogramme.

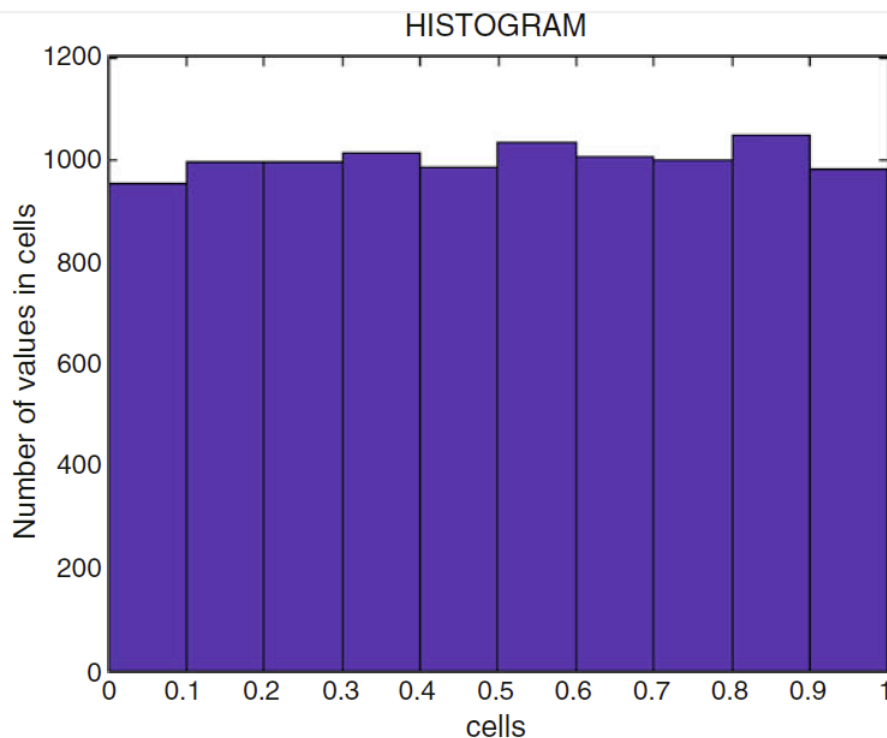


Fig.3 Histogramme, $M = 10$

Soit N_i le nombre de valeurs de la variable aléatoire X dans la $i^{\text{ème}}$ cellule. Ensuite, la probabilité pour que la variable aléatoire appartienne à la $i^{\text{ème}}$ cellule est approximée par une quantité N_i / N , appelé le rapport de fréquence,

$$P\{(i-1)\Delta x < x \leq i\Delta x\} \approx N_i / N = \text{hist}(X, M) / N \quad (2)$$

Le rapport de fréquence dans MATLAB est obtenu en divisant l'histogramme par N . Le tracé est représenté sur la figure 4.

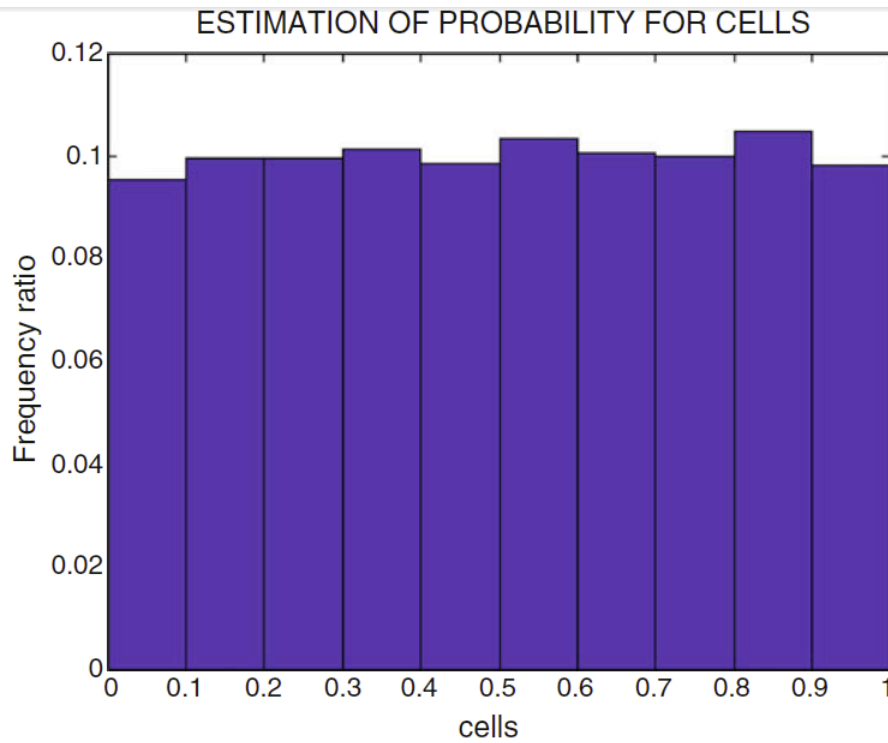


Fig.4 Estimation des probabilités

Ensuite, à partir de (1) et (2), on peut estimer la PDF dans la cellule comme suit :

$$f_x(i\Delta x) \approx \frac{P\{(i-1)\Delta x < x \leq i\Delta x\}}{\Delta x} = \frac{N_i}{N} \frac{1}{\Delta x} \quad (3)$$

D'après la formule (3), une estimation de la PDF au sein d'une cellule donnée est obtenue en divisant la probabilité pour que la variable aléatoire appartienne à la cellule (rapport de fréquence) par la longueur de la cellule Δx , qui est ici égal à $1/10=0.1$.

De (2) et (3), on obtient :

$$PDF \approx P\left\{X \text{ appartient à la cellule } i, i = 1, \dots, M\right\} / \Delta x = \frac{hist(x, M)}{N \cdot \Delta x} \quad (4)$$

L'estimation de la PDF est donnée à la figure 5.

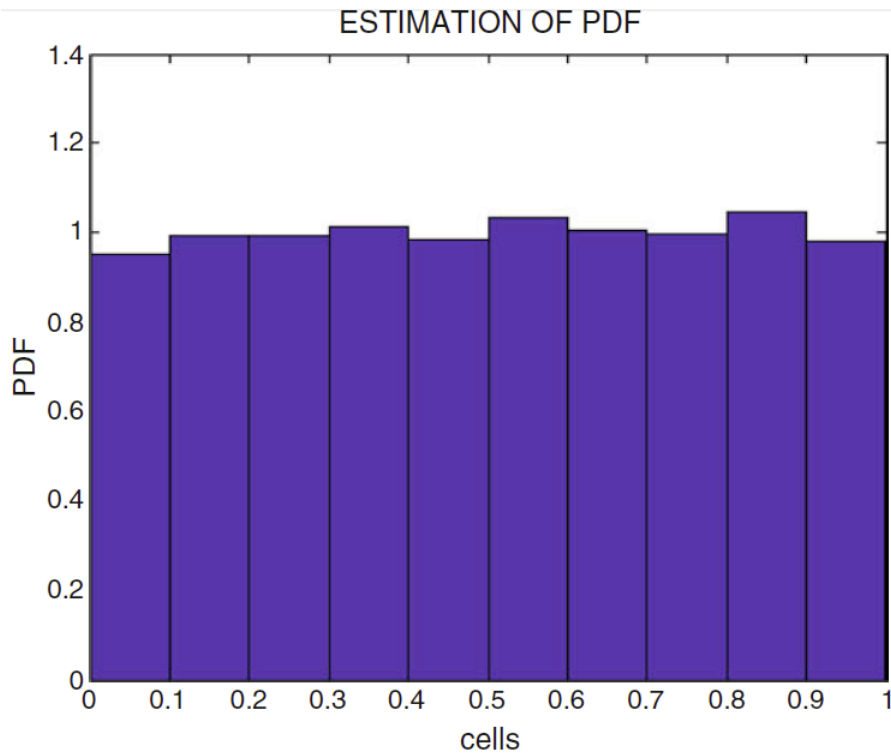


Fig.5 Estimation de la PDF

On notera que tous les tracés sur les figures 3, 4 et 5 ont la même forme et que la seule différence est la mise à l'échelle d'amplitude. Voilà pourquoi certains auteurs appellent l'histogramme le tracé d'estimation de la PDF. Cependant, nous considérons qu'il est important de clarifier la différence entre l'histogramme, les probabilités, et la PDF, parce que l'estimation de la PDF à partir de la Fig. 5 correspond au format mathématique de la PDF de la Fig. 1.a.

III.2. Fonction de distribution

Après, nous estimons la distribution à partir de la PDF en utilisant la formule suivante :

$$F_X(x) = \int_{-\infty}^x f_X(v) dv. \quad (5)$$

La fonction Matlab `cumsum.m` réalise cette estimation. La distribution estimée est donnée par la figure 6.

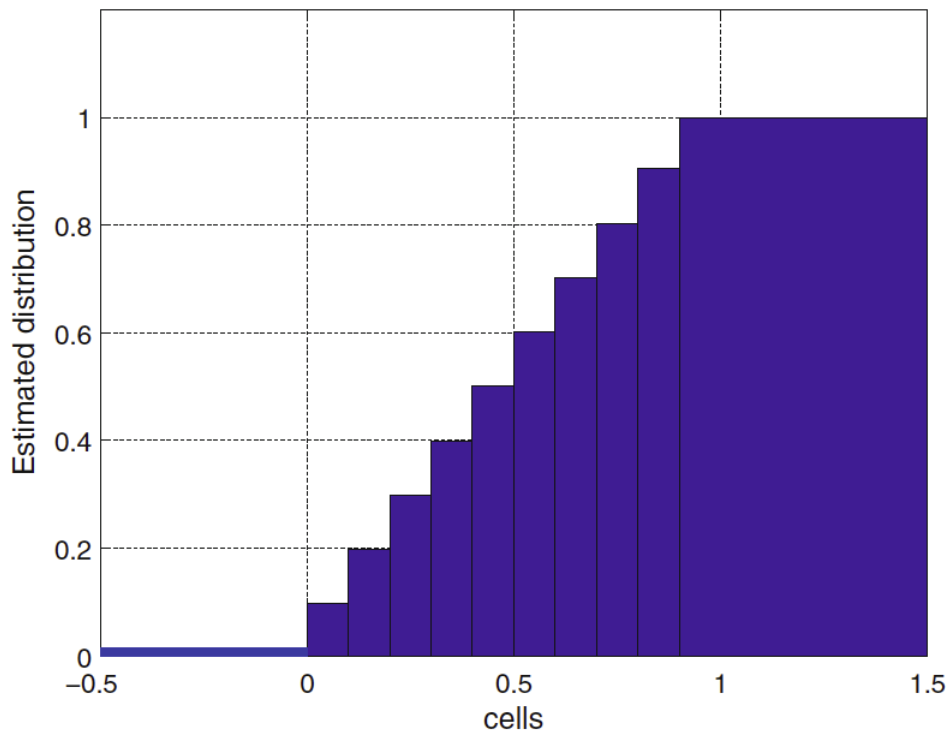


Fig.6 Distribution estimée

IV. Travail demandé

- 1- i) Générer puis tracer un vecteur ligne de 1000 valeurs aléatoires uniformes entre 0 et 1.
 ii) Tracer en rouge la partie du signal allant de la valeur 200 à la valeur 700.
 iii) Retracer le signal en changeant l'axe des ordonnées en [-1 2].
- 2- Montrer que la transformation linéaire d'une variable aléatoire uniforme résulte dans une variable uniforme, en utilisant à titre d'exemple, la variable aléatoire uniforme X dans l'intervalle [0, 1] et la transformation :

$$Y = - 2 X + 3$$
 Tracer X et Y pour 1500 valeurs aléatoires.
- 3- En utilisant le fichier MATLAB rand.m qui génère une variable aléatoire uniforme X dans l'intervalle [0, 1], écrire un programme MATLAB pour générer la variable uniforme dans l'intervalle désiré [R₁, R₂]. Considérons, par exemple, R₁ = - 2 et R₂ = 4. Tracer cette variable.
- 4- Estimer et tracer la PDF et la distribution de la variable aléatoire uniforme X dans l'intervalle [0.5, 3.5].
- 5- Générer une variable aléatoire uniforme dans l'intervalle [3, 7] et estimer sa PDF, sa valeur moyenne et sa variance.

TP 2 Processus aléatoires

1. Qu'est-ce qu'un processus aléatoire ?

Comme dans le cas d'une variable aléatoire, un processus aléatoire peut être défini à l'aide d'une expérience aléatoire. Rappelez-vous qu'une variable aléatoire est obtenue en attribuant des numéros sur l'axe x à tous les résultats possibles de s_i . Cependant, si au lieu des nombres, des fonctions temporelles $x(t, s_i)$ sont associées aux résultats de s_i à partir d'un espace d'échantillons S , on obtient un processus aléatoire $X(t, s)$, comme représenté sur la Fig.1.

De cette façon, le processus aléatoire $X(t, s)$ est une fonction qui affecte la famille des fonctions temporelles $x_i(t)$ à un espace d'événements S . En général, cette fonction peut être réelle ou complexe. Cependant, nous considérons ici que les processus réels.

L'ensemble de toutes les fonctions temporelles est appelé l'ensemble, et une fonction temporelle particulière est appelée une réalisation ou un membre d'un ensemble. En gardant à l'esprit que l'espace de cartographie S à un ensemble résulte d'un processus aléatoire, nous allons utiliser la dénotation $X(t)$ pour un processus et $x_i(t)$ pour une $i^{\text{ème}}$ réalisation particulière.

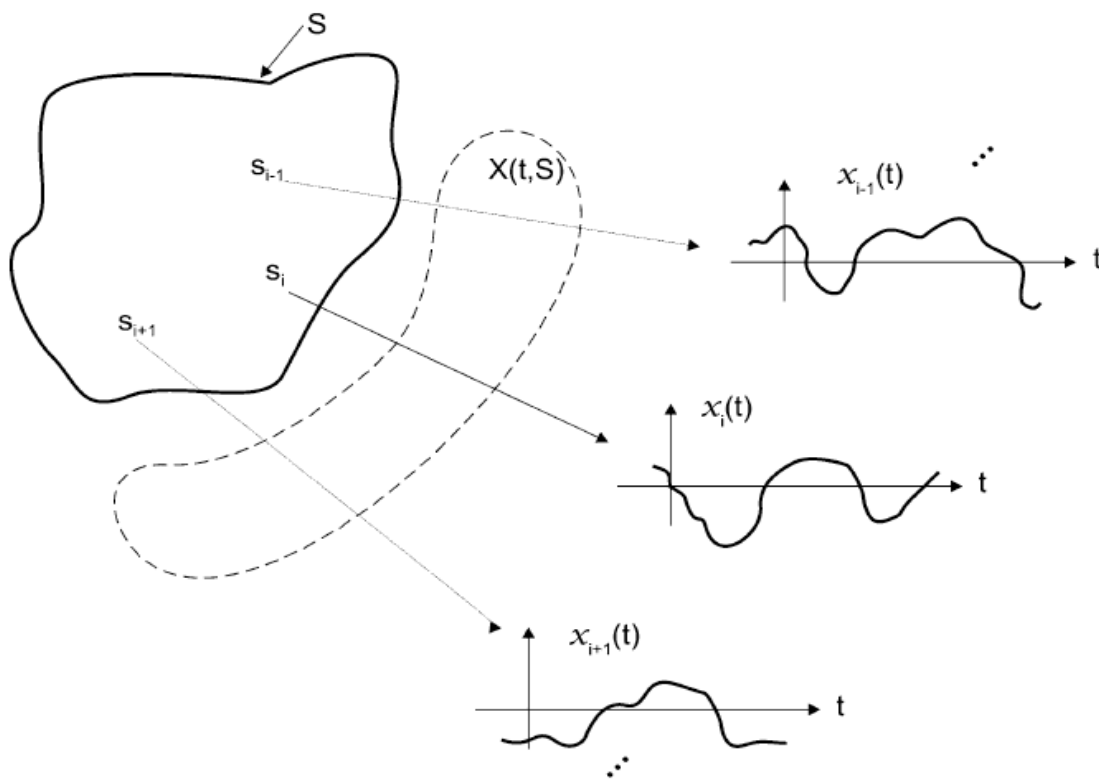


Fig.1 Mapping de l'espace des événements S vers l'espace (x, y)

1.1. Processus aléatoire déterministe et non déterministe

Un processus aléatoire est dit déterministe si toutes les valeurs futures de toute réalisation peuvent être déterminées à partir de ses valeurs précédentes. A titre d'exemple, considérons des réalisations sinusoïdales où l'amplitude varie pour les différentes réalisations, comme le montre la Fig. 2a. Les valeurs futures de chaque réalisation sont représentées avec des lignes en pointillés et sont complètement connues. Dans le cas contraire, représentée par la Fig. 2b., le processus est dit non déterministe. Les formes d'onde de réalisations sont irrégulières et ne peuvent pas être décrites en utilisant une expression mathématique. Par conséquent, les futures valeurs indiquées avec les lignes en pointillés ne peuvent pas être déterminées avec précision.

De la discussion précédente, notez que le terme «aléatoire» dans un processus aléatoire ne soit pas lié à une forme d'onde de ses réalisations, mais avec une incertitude dans laquelle la fonction de temps est affectée à un résultat particulier. De même, comme dans une expérience de lancement de dé, on sait avec certitude qu'un nombre 1 jusqu'à 6 apparaît, et l'incertitude est que ces chiffres apparaissent dans chaque lancé.

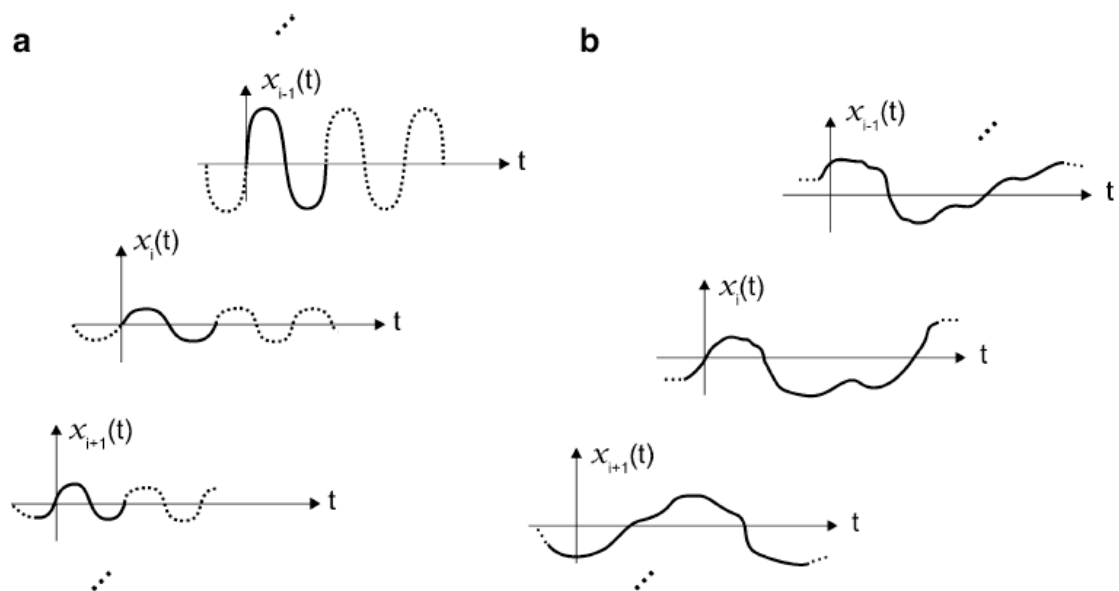


Fig. 2 Processus aléatoire déterministe et non déterministe

1.2. Processus aléatoire continu et discret

La classification des processus aléatoires peut être effectuée selon que les amplitudes et les temps sont des valeurs continues ou discrètes. Comme à la fois l'amplitude et le temps peuvent être soit continus soit discrets, toutes les combinaisons possibles sont présentées par la Fig. 3.

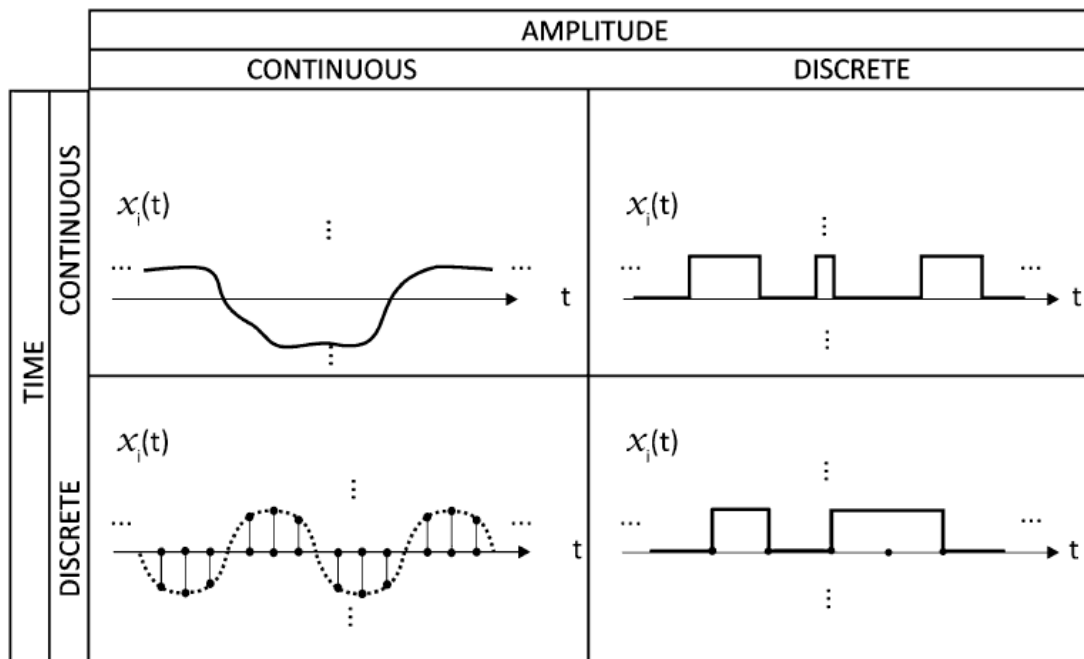


Fig.3 Processus aléatoire continu et discret

2. Statistiques d'un processus aléatoire

2.1. Description d'un processus en un point

Nous avons déjà mentionné qu'un processus aléatoire dans un temps déterminé t devient une variable aléatoire. Considérons l'instant de temps t_1 . Une variable aléatoire définie à un instant de temps t_1 est une variable aléatoire X_1 . La fonction de distribution de la variable, notée $F_{X_1}(x_1, t_1)$, est définie comme suit :

$$F_{X_1}(x_1, t_1) = P\{X_1 \leq x_1; t_1\} \quad (1)$$

Notons que la définition (1) est la même que la définition de la distribution d'une variable aléatoire, la seule différence étant que la répartition (1) dépend de l'instant de temps t_1 .

La signification de la distribution (1) est expliquée par la Fig. 4. Pendant l'instant de temps spécifié t_1 et la valeur x_1 , la distribution (1) présente la probabilité que,

à l'instant t_1 toutes les réalisations sont inférieures ou égales à une valeur x_1 . Cette probabilité peut également être interprétée en utilisant un rapport de fréquence, comme il est décrit ci-dessous.

Considérons qu'une expérience aléatoire est exécutée n fois et une réalisation particulière est donnée à chaque résultat, comme représenté par la Fig. 4. Soit $n(x_1, t_1)$ le nombre total de succès où les amplitudes des réalisations à l'instant de temps t_1 ne sont pas plus que x_1 .

Prenant n très grand, la probabilité souhaitée peut être approximée comme :

$$F_{x_1}(x_1, t_1) \approx \frac{n(x_1, t_1)}{n} \quad (2)$$

La distribution (2) est appelée une distribution du premier ordre, ou une distribution unidimensionnelle d'un processus $X(t)$. Notons que la distribution unidimensionnelle est obtenue en observant un processus en un instant dans le temps.

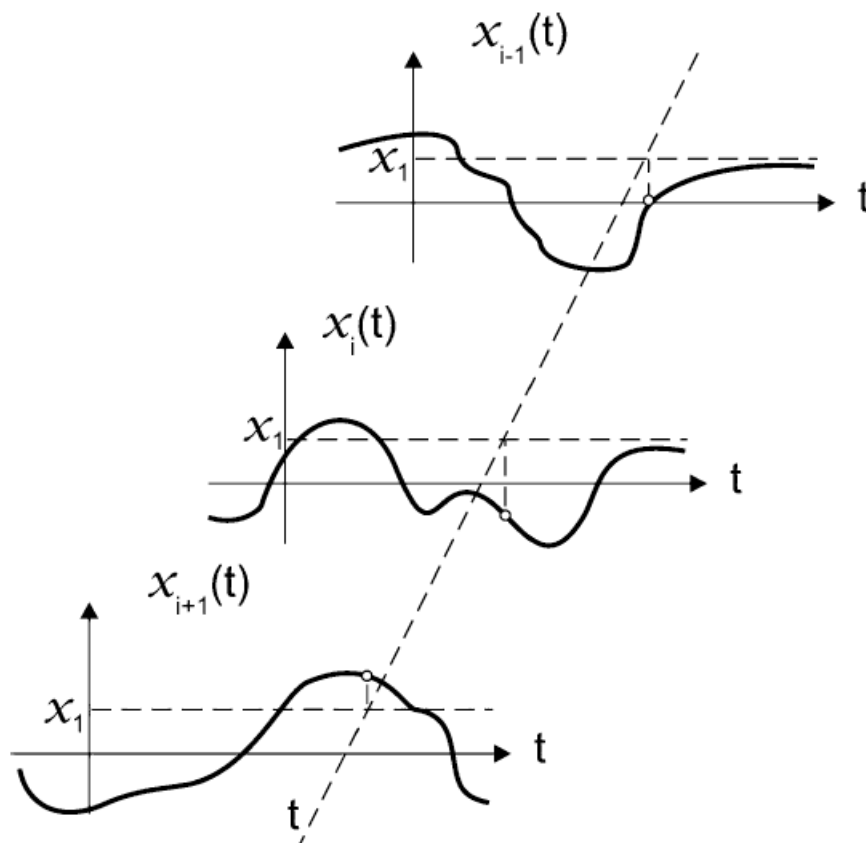


Fig.4 Description d'un processus en un instant de temps particulier

Une fonction de densité est obtenue comme la dérivation de la fonction de distribution correspondante,

$$f_{X_1}(x_1, t_1) = \frac{\partial F_{X_1}(x_1; t_1)}{\partial x_1} = \frac{P\{x_1 < X_1 \leq x_1 + dx_1; t_1\}}{dx_1} \quad (3)$$

La fonction de densité (3) est appelée fonction de densité du premier ordre ou une densité unidimensionnelle.

2.2. Description d'un processus en deux points

Nous observons le processus en deux points notés t_1 et t_2 . Les variables aléatoires correspondantes sont X_1 et X_2 . Une fonction de distribution conjointe, une distribution de second ordre ou une distribution à deux dimensions est définie comme :

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2; t_1, t_2) &= P\{X(t_1) \leq x_1, X(t_2) \leq x_2\} \\ &= P\{X_1 \leq x_1, X_2 \leq x_2; t_1, t_2\} \end{aligned} \quad (4)$$

Notons aussi que cette définition est la même que celle de la distribution conjointe de deux variables à la seule exception que maintenant la distribution conjointe dépend des instants de temps t_1 et t_2 .

La signification de cette distribution peut être interprétée en utilisant le rapport de fréquence comme le montre la Fig. 5. Considérons un nombre total de sorties pour lesquelles les réalisations correspondantes ne sont pas plus que x_1 en un instant de temps t_1 et ne sont pas plus que x_2 en un instant de temps t_2 . Le nombre obtenu est dénoté comme $n(x_1, x_2, t_1, t_2)$. Ensuite, supposons que n est grand, la fonction de distribution peut être exprimée comme :

$$F_{X_1, X_2}(x_1, x_2; t_1, t_2) = P\{X_1 \leq x_1, X_2 \leq x_2; t_1, t_2\} \approx \frac{n(x_1, x_2; t_1, t_2)}{n} \quad (5)$$

La fonction de densité correspondante est donnée par :

$$\begin{aligned} f_{X_1, X_2}(x_1, x_2; t_1, t_2) &= \frac{\partial^2 F_{X_1, X_2}(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2} \\ &= \frac{P\{x_1 < X(t_1) \leq x_1 + dx_1, x_2 < X(t_2) \leq x_2 + dx_2; t_1, t_2\}}{dx_1 dx_2} \\ &= \frac{P\{x_1 < X_1 \leq x_1 + dx_1, x_2 < X_2 \leq x_2 + dx_2; t_1, t_2\}}{dx_1 dx_2} \end{aligned} \quad (6)$$

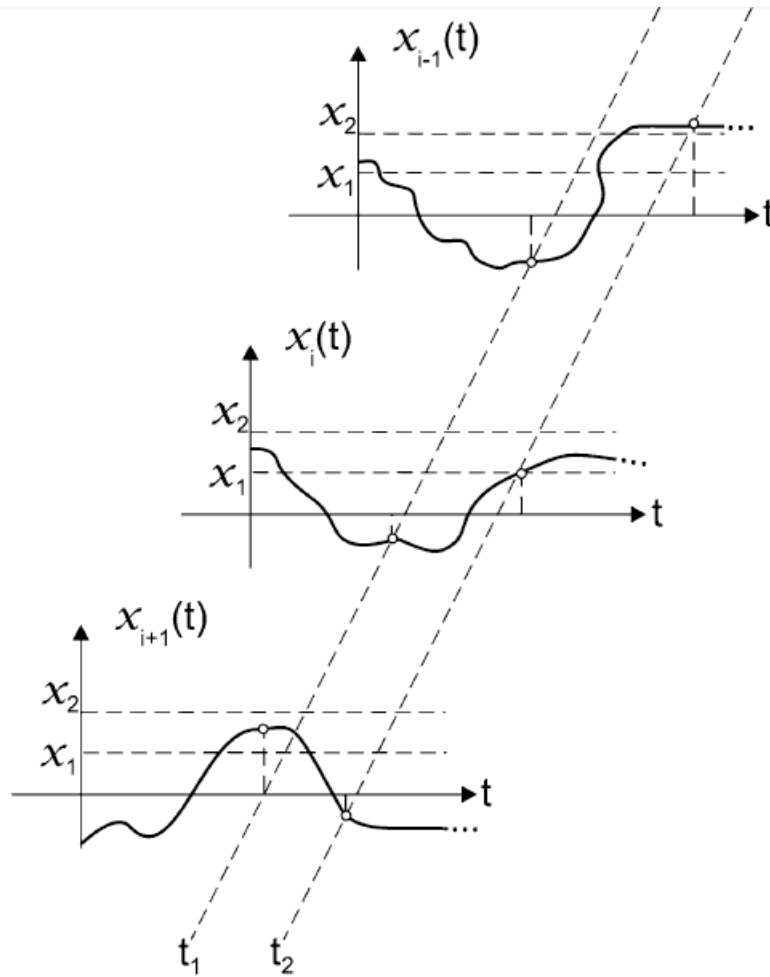


Fig.5 Description d'un processus en deux instants de temps

3. Processus aléatoire stationnaire

Un processus est appelé stationnaire si aucune de ses caractéristiques statistiques ne changent avec le temps.

Nous pouvons définir différents types de processus stationnaires en fonction des caractéristiques statistiques observées.

Un processus est dit stationnaire du premier ordre si sa première fonction de densité est indépendante d'un décalage temporel Δ , c'est-à-dire :

$$f_X(x,t) = f_X(x,t + \Delta) \quad (7)$$

pour tout t et Δ .

Donc,

$$f_X(x,t) = f_X(x) \quad (8)$$

De même, un processus est dit stationnaire du second ordre si sa fonction de densité de seconde ordre a la propriété suivante pour toute valeur de t_1, t_2 , et Δ :

$$f_{X_1, X_2}(x_1, x_2; t_1, t_2) = f_{X_1, X_2}(x_1, x_2; t_1 + \Delta, t_2 + \Delta) \quad (9)$$

On notera que, dans ce cas, la fonction de densité conjointe (9) est dépendante d'une différence entre les instants de temps t_2 et t_1 ,

$$\tau = t_2 - t_1 \quad (10)$$

et peut donc être exprimée comme suit :

$$f_{X_1, X_2}(x_1, x_2; t_1, t_2) = f_{X_1, X_2}(x_1, x_2; t_1 + \Delta, t_2 + \Delta) = f_{X_1, X_2}(x_1, x_2; \tau) \quad (11)$$

En généralisant la discussion précédente à un cas de n variables aléatoires, X_i , $i = 1, \dots, N$, nous pouvons définir un processus stationnaire d'ordre n . Dans ce cas, une fonction de densité conjointe de dimension n a la propriété suivante, pour toutes les valeurs de t_1, \dots, t_n et Δ :

$$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n; t_1 + \Delta, t_2 + \Delta, \dots, t_n + \Delta) \quad (12)$$

Un processus stationnaire d'ordre n est aussi un processus stationnaire d'ordre k , où $k = 1, \dots, n-1$.

4. Valeur moyenne

Une valeur moyenne pour une seule variable peut également être introduite pour un processus observé à un instant particulier, comme le montre la Fig. 6. Le processus à un instant de temps t_1 est une variable aléatoire $X_1(t_1) = X_1$. Soit la plage continue de la variable X_1 divisée en k éléments Δx , tels qu'ils sont si petits que si la variable est dans l'intervalle Δx , nous disons qu'il est égal à Δx . De même, si la variable est dans le $i^{\text{ème}}$ intervalle $i\Delta x$, nous disons qu'il est égal à $i\Delta x$. De même, comme dans le cas d'une seule variable, on trouve d'abord une valeur moyenne empirique comme résultat d'une expérience réalisée.

L'expérience est répétée N fois dans les mêmes conditions. Le processus $X(t)$ est obtenu en assignant N réalisations $x(t)$ à la sortie de chaque expérience. Considérons toutes les réalisations à l'instant de temps t_1 , et supposons que nous savons ce qui suit :

N_1 réalisations sont dans l'intervalle $[0, \Delta x]$ (comme la réalisation $x_{n+1}(t)$ de la figure 6)

N_2 réalisations sont dans l'intervalle $[\Delta x, 2\Delta x]$ (comme la réalisation $x_{n-1}(t)$ de la figure 6)

...

N_i réalisations sont dans l'intervalle $[(i-1)\Delta x, i\Delta x]$ (comme la réalisation $x_n(t)$ de la figure 6)

...

N_k réalisations sont dans l'intervalle $[(k-1)\Delta x, k\Delta x]$

Une valeur moyenne empirique du processus à l'instant de temps t_1 est égale à :

$$\begin{aligned} X_{emp.av}(t_1) &= \frac{N_1\Delta x + N_2 2\Delta x + \dots + N_i i\Delta x + \dots + N_k k\Delta x}{N} \\ &= \sum_{i=1}^k \frac{N_i}{N} i\Delta x = \sum_{i=1}^k \left(\frac{N_i}{N} \frac{1}{\Delta x} \right) i\Delta x \Delta x \end{aligned} \quad (13)$$

Parce que N est très grand et Δx très faible, une valeur moyenne empirique devient indépendante d'une expérience et elle se rapproche de la valeur moyenne d'un processus :

$$X_{emp.av}(t_1) \rightarrow \overline{X(t_1)} = m(t_1) = \int_{-\infty}^{\infty} x f_X(x; t) dx, \text{ comme } N \rightarrow \infty, \Delta x \rightarrow dx, i\Delta x \rightarrow x \quad (14)$$

A un autre instant de temps t_2 , il est possible d'obtenir une valeur moyenne différente :

$$m(t_2) = \int_{-\infty}^{\infty} x f_X(x; t_2) dx \quad (15)$$

En général, une valeur moyenne du processus dépend du temps :

$$m(t) = \overline{X(t)} = E\{X(t)\} = \int_{-\infty}^{\infty} x f_X(x; t) dx \quad (16)$$

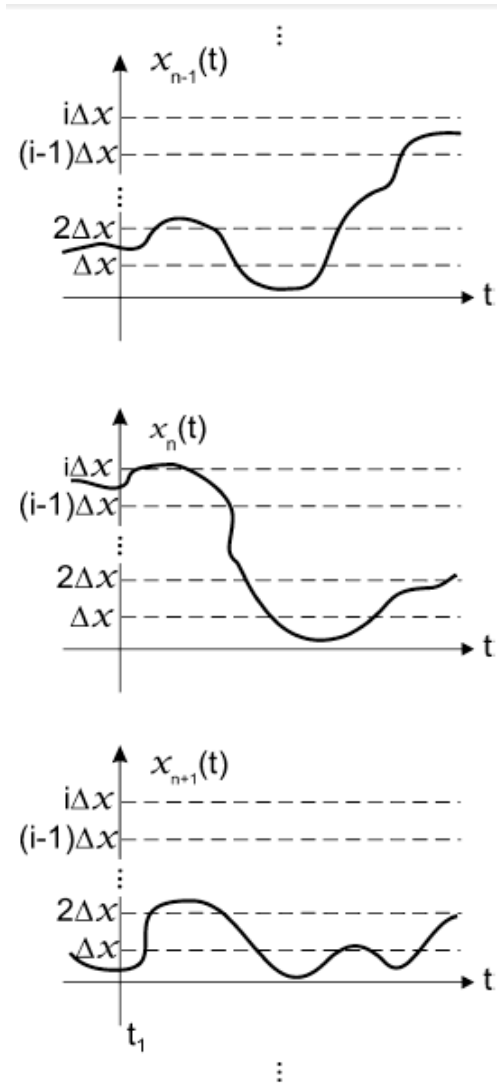


Fig.6 Explication de la valeur moyenne du processus

5. Fonction d'autocorrélation

5.1. Définition

L'analyse statistique d'un processus aléatoire est effectuée en utilisant les caractéristiques qui caractérisent l'ensemble de son processus. L'une des caractéristiques les plus importantes est la fonction d'autocorrélation.

Une fonction d'autocorrélation est définie comme une valeur possible des variables aléatoires du processus défini aux instants t_1 et t_2 ,

$$R_{XX}(t_1, t_2) = \overline{X(t_1)X(t_2)} = \overline{X_1 X_2} \quad (17)$$

ou de manière équivalente :

$$R_{XX}(t_1, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X_1 X_2}(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (18)$$

Notons qu'une fonction d'autocorrélation est généralement dépendante des instants de temps t_1 et t_2 . Dans de nombreux problèmes pratiques, des solutions peuvent être simplifiées si nous pouvons supposer qu'une fonction d'autocorrélation ne dépend pas des instants t_1 et t_2 , mais plutôt de leur différence $\tau = t_2 - t_1$.

5.2. Processus stationnaire au sens large (Wide Sense stationary process)

Un processus aléatoire est dit stationnaire au sens large si les deux conditions suivantes sont satisfaites : la valeur moyenne est constante et la fonction d'autocorrélation dépend seulement de la différence de temps $\tau = t_2 - t_1$.

$$(a) \quad \overline{X(t)} = \text{const} \quad (19)$$

$$(b) \quad R_{XX}(\tau) = \overline{X(t)X(t+\tau)} = \int_{-\infty}^{\infty} x_1 x_2 f_{X_1 X_2}(x_1, x_2; \tau) dx_1 dx_2 \quad (20)$$

6. Processus ergodique

6.1. Moyenne Temps (Time Averaging)

Nous considérons des valeurs moyennes de temps obtenues en faisant la moyenne des réalisations particulières du processus dans le temps. De même, nous avons introduit la notation $E\{\}$ pour la moyenne d'ensemble, nous introduisons la notation $A\{\}$ pour le calcul de la moyenne du temps.

Une moyenne de temps d'une fonction $g(x(t))$ est définie comme suit :

$$A\{g(x(t))\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} g(x(t)) dt \quad (21)$$

Deux valeurs moyennes de temps : la valeur moyenne du temps et de la fonction d'autocorrélation de temps ont une importance particulière. Ces valeurs moyennes sont obtenues en observant une réalisation $x(t)$ du processus dans le temps.

Une valeur moyenne de temps est définie comme suit :

$$A\{x(t)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt = \overline{x(t)} \quad (22)$$

De la même façon, on peut définir la fonction d'autocorrélation temporelle pour une réalisation $x(t)$ comme suit :

$$R_{xx}(\tau) = A \{x(t)x(t+\tau)\} = \overline{x(t)x(t+\tau)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t+\tau) dt \quad (23)$$

où $\overline{x(t)}$ représente la moyenne d'une réalisation $x(t)$ du processus $X(t)$ dans le temps.

6.2. Qu'est-ce que l'ergodicité ?

Un processus est dit ergodique si toutes ses moyennes statistiques sont égales à ses moyennes de temps correspondantes. De (22), il est clair que pour un processus ergodique, toute moyenne de temps ne peut pas être une fonction du temps. En d'autres termes, dans les processus ergodiques, les moyennes statistiques correspondantes ne peuvent pas être fonction du temps. Sachant que seuls les processus stationnaires possèdent ces caractéristiques, nous pouvons conclure que le processus ergodique doit être stationnaire. Cependant, le contraire est faux (par exemple, il existe des processus stationnaires qui ne sont pas des processus ergodiques). A titre d'exemple, considérons le procédé de tensions continues, représentées sur la Fig.7. Ce processus est évidemment stationnaire parce que les caractéristiques statistiques sont indépendantes des instants de temps dans lesquels le processus est observé. Cependant, parce que les amplitudes des réalisations sont différentes, les moyennes de temps varient d'une réalisation à l'autre. Par conséquent, le processus est stationnaire mais non ergodique.

Les relations mutuelles entre stationnarité et ergodicité peuvent être illustrées, comme le montre Fig.8. L'ergodicité permet de décrire un processus en connaissant une seule réalisation.

Ceci est d'un intérêt pratique parce que généralement, on a connaissance d'une seule fonction de temps d'un processus et, sur la base de l'ergodicité, peut trouver une moyenne statistique du processus en trouvant le temps moyen d'une seule réalisation du processus.

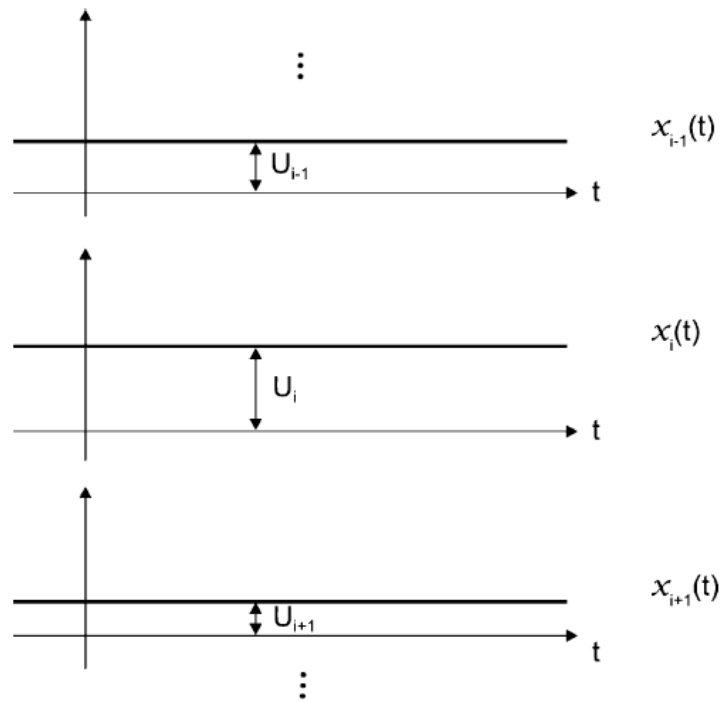


Fig.7 Exemple d'un processus stationnaire qui n'est pas ergodique

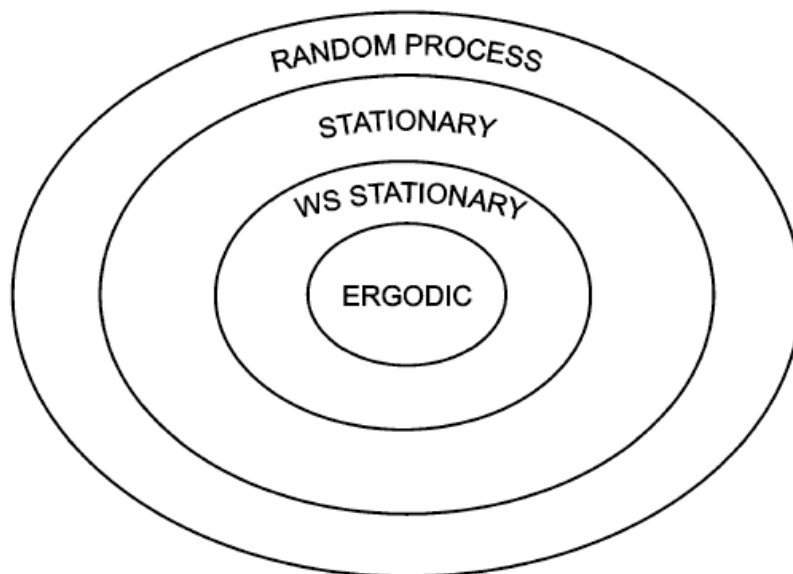


Fig. 8 Stationnarité et ergodicité

Travail demandé

Exercice 1 : Générer un processus aléatoire uniforme avec $N = 2000$ réalisations où chaque réalisation contient $M = 2000$ éléments. Le processus est uniforme dans l'intervalle $[-1, 1]$.

- 1) Observer le processus en trois points différents : $M_1=1$, $M_2=530$ et $M_3=800$.
- 2) Désigner les variables aléatoires correspondantes par X_1 , X_2 et X_3 et donner leurs tracés.
- 3) Estimer les valeurs moyennes des variables X_1 , X_2 et X_3 .
- 4) Estimer les variances des variables X_1 , X_2 et X_3 .
- 5) Tracer les histogrammes des variables X_1 , X_2 et X_3 .
- 6) Estimer les PDF des variables X_1 , X_2 et X_3 .
- 7) Est-ce processus stationnaire ?

Exercice 2 : Générer un processus aléatoire normal avec $N = 1000$ réalisations où chaque réalisation contient $M = 10000$ éléments. La moyenne et la variance sont respectivement 0 et 1.

- 1) Observer le processus en trois points différents : $M_1=91$, $M_2=780$ et $M_3=900$.
- 2) Désigner les variables aléatoires correspondantes par X_1 , X_2 et X_3 et donner leurs tracés.
- 3) Estimer les valeurs moyennes des variables X_1 , X_2 et X_3 .
- 4) Estimer les variances des variables X_1 , X_2 et X_3 .
- 5) Tracer les histogrammes des variables X_1 , X_2 et X_3 .
- 6) Estimer les PDF des variables X_1 , X_2 et X_3 .
- 7) Est-ce processus stationnaire ?

Exercice 3 : Générer un processus aléatoire uniforme avec $N = 3000$ réalisations où chaque réalisation contient $M = 3000$ éléments. Le processus est uniforme dans l'intervalle $[1, 5]$.

Observer le processus à un point $M_1 = 1203$. La variable aléatoire X_1 est obtenue. Choisissez une réalisation du processus $N_1 = 547$.

Les amplitudes de la réalisation sont la variable aléatoire Y_1 .

$M_1 = 1203; N_1 = 547 :$

- 1) Tracer le processus au point M_1 . Tracer la N_1^{th} réalisation.
- 2) Estimer les valeurs moyennes des variables X_1 et Y_1 .
- 3) Estimer les variances des variables X_1 et Y_1 .
- 4) Tracer les histogrammes des variables X_1 et Y_1 .
- 5) Estimer les PDF des variables X_1 et Y_1 .
- 6) Est-ce processus ergodique ?

TP 3 Densité spectrale de puissance

But : Le travail pratique 2 a été focalisé sur la description d'un processus aléatoire dans le temps. Cependant, comme dans la description des signaux déterministes, il est intéressant de décrire également un procédé aléatoire dans le domaine des fréquences. Nous considérerons ici que les processus réels.

1. Transformation de Fourier des signaux déterministes

La description spectrale d'un signal déterministe $f(t)$ est obtenue en prenant la transformée de Fourier $F(\omega)$, également connu sous le spectre d'un signal $f(t)$:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (1)$$

Si $f(t)$ est la tension, alors $F(\omega)$ représente la tension par Hertz et décrit comment la tension est répartie en fréquences.

Pour que la transformation de Fourier (1) existe, la condition suivante doit être satisfaite :

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty \quad (2)$$

La transformation de Fourier inverse trouve le signal $f(t)$ dans le domaine temporel à partir de son spectre

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} d\omega \quad (3)$$

La question que l'on pourrait naturellement se poser est : "Peut-on appliquer également la transformation de Fourier aux signaux aléatoires ?"

La réponse est fournie dans la section suivante.

2. Comment appliquer la transformation de Fourier aux signaux aléatoires?

L'application directe de la transformée de Fourier à un processus aléatoire est impossible pour les raisons énumérées ci-dessous :

- Il y a un problème de l'existence de la transformée de Fourier (2), à savoir, pour la majorité des réalisations du processus, la condition (2) n'est pas satisfaite et la transformée de Fourier peut ne pas exister.
- Même dans un cas dans lequel la transformée de Fourier d'une réalisation particulière existe, il est évident que le résultat ne peut pas représenter l'ensemble du processus. La transformée de Fourier des réalisations individuelles serait différente.
- Dans la majorité des cas, les réalisations d'un processus ont des formes irrégulières et ne peuvent pas être représentées sous formes analytiques et, par conséquent, il y a un problème de calcul pratique de la transformée de Fourier.

Ensuite, nous allons décrire comment surmonter ces difficultés en appliquant la transformée de Fourier à un processus aléatoire.

2.1. Problème de l'existence de la transformation de Fourier

Pour une réalisation particulière $x(t)$ d'un processus aléatoire $X(t)$, qui existe dans l'intervalle $[-\infty, \infty]$, la condition

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty \quad (4)$$

ne peut pas être satisfaite et, par conséquent, la transformée de Fourier n'existe pas. L'énergie moyenne W d'une réalisation particulière serait également infinie

$$W = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} x^2(t) dt \quad (5)$$

Cependant, une puissance moyenne P

$$P = \lim_{T \rightarrow \infty} \frac{W}{T} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t) dt \quad (6)$$

est finie. Cette conclusion suggère l'application de la transformée de Fourier à une puissance plutôt qu'à l'amplitude d'une réalisation. Cette idée est développée ci-dessous.

Prenant en considération la part d'une réalisation $x(t)$ d'un processus $X(t)$ dans l'intervalle $[-T/2, T/2]$, comme représenté sur la Fig. 1.

$$x_T(t) = \begin{cases} x(t) & \text{pour } -T/2 \leq t \leq T/2 \\ 0 & \text{ailleurs} \end{cases} \quad (7)$$

Pour l'intervalle fini T , x_T doit satisfaire à la condition (4) et, par conséquent, x_T aura la transformée de Fourier :

$$X_T(\omega) = \int_{-T/2}^{T/2} x_T(t) e^{-j\omega t} dt = \int_{-\infty}^{\infty} x_T(t) e^{-j\omega t} dt = \int_{-T/2}^{T/2} x(t) e^{-j\omega t} dt \quad (8)$$

L'énergie du signal $X_T(t)$ est la suivante :

$$W_T = \int_{-T/2}^{T/2} x_T^2(t) dt = \int_{-\infty}^{\infty} x_T^2(t) dt \quad (9)$$

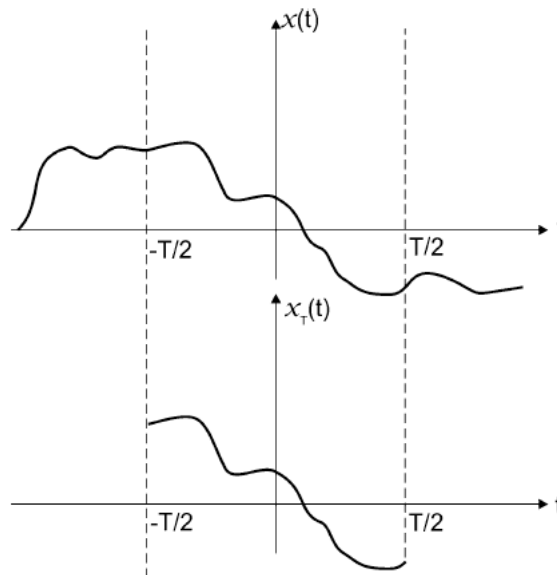


Fig.1 Réalisation $x(t)$ dans l'intervalle $[-T/2, T/2]$

Selon le théorème de Parseval, l'énergie W_T peut être exprimée comme suit :

$$W_T = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X_T(\omega)|^2 d\omega \quad (10)$$

L'expression obtenue est également l'énergie d'un signal $x(t)$ dans l'intervalle $[T/2, T/2]$. Par conséquent, l'énergie de l'ensemble du signal $x(t)$ est obtenu à partir de (10), lorsque $T \rightarrow \infty$,

$$W = \lim_{T \rightarrow \infty} W_T \quad (11)$$

A partir de (11), la puissance du signal $x(t)$ est la suivante :

$$P_{xx} = \lim_{T \rightarrow \infty} \frac{W_T}{T} \quad (12)$$

En utilisant (10), nous sommes finalement arrivés à :

$$P_{xx} = \lim_{T \rightarrow \infty} \frac{1}{T} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} |X_T(\omega)|^2 d\omega \right] \quad (13)$$

Cette expression peut être réécrite comme suit :

$$P_{xx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{|X_T(\omega)|^2}{T} d\omega \quad (14)$$

Le membre gauche de l'équation (14) est une puissance. Par conséquent, l'expression dans l'intégrale $\lim_{T \rightarrow \infty} |X_T(\omega)|^2 / T$, est une densité spectrale de puissance (DSP) d'une unité [Puissance/Hertz].

La DSP est désignée par $S_{xx}(\omega)$

$$S_{xx}(\omega) = \lim_{T \rightarrow \infty} \frac{|X_T(\omega)|^2}{T} \quad (15)$$

En utilisant (14) et (15), nous avons :

$$P_{xx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) d\omega \quad (16)$$

A partir de (15), nous pouvons conclure que la DSP obtenue existe et présente le contenu spectral d'une réalisation particulière d'un processus $X(t)$.

2.2. Comment pouvons-nous obtenir la DSP d'un processus à partir de la DSP d'une réalisation particulière ?

Afin d'obtenir la densité spectrale de puissance $S_{xx}(\omega)$ d'un processus, nous devons trouver la valeur moyenne des densités spectrales de puissance sur tout l'ensemble. En d'autres termes, nous devons trouver une valeur moyenne de puissance spectrale des densités $S_{xx}(\omega)$ pour toutes les réalisations :

$$S_{xx}(\omega) = E\{S_{xx}(\omega)\} \quad (17)$$

Par conséquent, nous avons :

$$S_{xx}(\omega) = E\left\{\lim_{T \rightarrow \infty} \frac{1}{T} |X_T(\omega)|^2\right\} = \lim_{T \rightarrow \infty} \frac{E\{|X_T(\omega)|^2\}}{T} \quad (18)$$

La puissance moyenne du processus est donnée ci-dessous :

$$P_{xx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) d\omega \quad (19)$$

2.3. Comment pouvons-nous trouver la densité spectrale (18) si les réalisations ont une forme irrégulière et ne peuvent être représentées sous forme analytique?

Il est évident que, dans ce cas, il existe un problème lié au calcul pratique de la DSP. La solution est donnée dans ce qui suit.

3. DSP et fonction d'autocorrélation

Comme mentionné précédemment, l'expression obtenue (18) pour la DSP ne convient pas pour une application pratique, car il est nécessaire de calculer la transformée de Fourier des amplitudes au carré d'un processus aléatoire, ce qui ne peut généralement pas être exprimé sous une forme analytique. Par

conséquent, il est intéressant d'exprimer la DSP sous une forme plus commode, comme cela est expliqué ci-après.

A partir de la transformée de Fourier d'un signal $x_T(t)$, définie en (8) :

$$X_T(\omega) = \int_{-\infty}^{\infty} x_T(t) e^{-j\omega t} dt = \int_{-T/2}^{T/2} x(t) e^{-j\omega t} dt \quad (20)$$

Nous écrivons :

$$\begin{aligned} |X_T(\omega)|^2 &= X_T(-\omega) X_T(\omega) \\ &= \int_{-T/2}^{T/2} x(t_1) e^{j\omega t_1} dt_1 \int_{-T/2}^{T/2} x(t_2) e^{-j\omega t_2} dt_2 \\ &= \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} x(t_1) x(t_2) e^{-j\omega(t_2-t_1)} dt_1 dt_2 \end{aligned} \quad (21)$$

En utilisant (18) et (21), et sachant que l'espérance est appliquée à un processus $X(t)$, nous avons :

$$S_{XX}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \overline{|X_T(\omega)|^2} = \lim_{T \rightarrow \infty} \frac{1}{T} \overline{\int_{-T/2}^{T/2} \int_{-T/2}^{T/2} X(t_1) X(t_2) e^{-j\omega(t_2-t_1)} dt_1 dt_2} \quad (22)$$

En interchangeant les opérations d'espérance et d'intégration, nous avons :

$$S_{XX}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} \overline{X(t_1) X(t_2)} e^{-j\omega(t_2-t_1)} dt_1 dt_2 \quad (23)$$

La valeur moyenne dans (23) présente une fonction d'autocorrélation :

$$\overline{X(t_1) X(t_2)} = R_{XX}(t_1, t_2) \quad (24)$$

résultante en :

$$S_{XX}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} R_{XX}(t_1, t_2) e^{-j\omega(t_2-t_1)} dt_1 dt_2 \quad (25)$$

En exprimant t_1 et t_2 comme t et $t + \tau$ respectivement (où $\tau = t_2 - t_1$), nous obtenons :

$$S_{XX}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2-t}^{T/2-t} \int_{-T/2}^{T/2} R_{XX}(t, t + \tau) e^{-j\omega\tau} dt d\tau \quad (26)$$

ou bien

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_{XX}(t, t + \tau) dt \right\} e^{-j\omega\tau} d\tau \quad (27)$$

L'expression entre les accolades présente une moyenne temporelle de la fonction d'autocorrélation du processus X(t)

$$A\{R_{XX}(t, t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_{XX}(t, t + \tau) dt \quad (28)$$

Si l'on suppose que le processus est au moins stationnaire au sens large, sa fonction d'autocorrélation ne dépend pas de temps absolu, mais seulement d'une différence de temps t, ce qui entraîne :

$$A\{R_{XX}(t, t + \tau)\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_{XX}(\tau) dt = R_{XX}(\tau) \quad (29)$$

En remplaçant (29) dans (27), on obtient finalement :

$$S_{XX}(\omega) = \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-j\omega\tau} d\tau \quad (30)$$

ou de manière équivalente

$$R_{XX}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{XX}(\omega) e^{j\omega\tau} d\tau \quad (31)$$

Les expressions obtenues (30) et (31) sont connues comme le théorème de Wiener-Khintchine, qui stipule qu'une DSP et une fonction d'autocorrélation d'un processus, qui est au moins stationnaire au sens large, sont la paire de transformée de Fourier.

L'importance de ce théorème est que l'on peut obtenir une caractéristique spectrale (par exemple, la DSP d'un processus aléatoire comme la transformée

de Fourier d'une fonction d'autocorrélation), qui est lui-même une fonction déterministe qui représente le processus.

Par conséquent, le théorème de Wiener-Khinchin résout tous les problèmes liés à l'application de la transformée de Fourier à un processus aléatoire, mentionné dans la section 2.

4. Propriétés de la DSP

La DSP a un certain nombre de propriétés utiles. Certaines d'entre elles sont présentées ci-dessous. Nous allons considérer seulement un processus réel qui est au moins stationnaire au sens large.

P.1 La densité de puissance est une fonction réelle.

P.2 La DSP est une fonction paire en pulsation ω

P.3 La puissance moyenne d'un processus aléatoire est obtenue comme la surface en dessous de la DSP, comme représenté sur la Fig. 2.

P.4 La DSP est une fonction non négative.

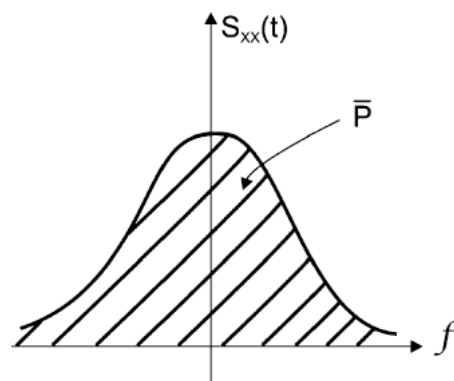


Fig. 2 Puissance moyenne et la densité spectrale

Travail demandé

- 1) En utilisant la méthode de périodogramme, estimer les DSPs et montrer leurs tracés pour deux signaux blanc X_1 (normal) et X_2 (uniforme), ayant $m = 0$ et $\sigma^2 = 1$.
- 2) i) Générer deux signaux non corrélés X_1 (normal) et X_2 (uniforme) ayant $m = 0$ et $\sigma^2 = 1$. Tracer les fonctions d'autocorrélation.
ii) Les signaux sinusoïdaux d'amplitude A et de fréquences $f_1=150\text{Hz}$ et $f_2=250\text{Hz}$ sont ajoutés respectivement aux signaux aléatoires X_1 et X_2 précédents. En utilisant la méthode de périodogramme, estimer leurs DSPs et montrer les tracés DSP pour $A = 1$ et $A = 2$.
- 3) En utilisant la méthode de périodogramme, estimer les DSPs et montrer leurs tracés pour les signaux blancs normaux avec :
 1. $m = 4$, et $\sigma^2 = 9$.
 2. $m = 5$ et $\sigma^2 = 16$.
- 4) En utilisant la méthode de périodogramme, estimer la DSP et montrer leurs tracés pour des signaux sinusoïdaux avec $A_1 = 3$, $f_1 = 270 \text{ Hz}$ ainsi que $A_2 = 8$, $f_2 = 175 \text{ Hz}$ dans un bruit blanc normal avec $m = 3$ et $\sigma^2 = 25$, comme représenté sur la Fig. 3.

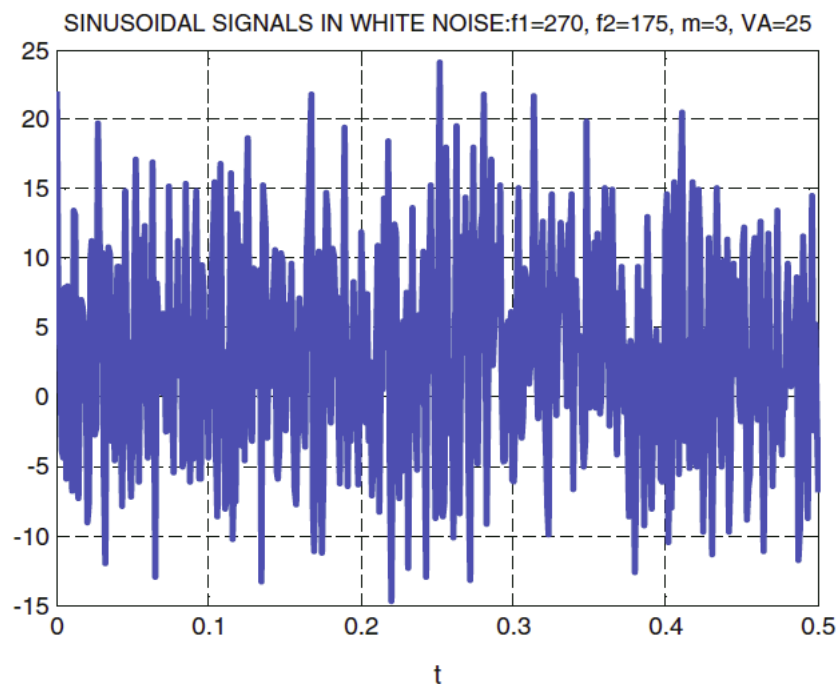


Fig. 3 Signaux sinusoïdaux en bruit blanc à moyenne non nulle

- 5) Étant donné un bruit normal avec $m = 4$ et $\sigma^2 = 4$.
- i) Si le bruit est filtré avec un filtre passe-bas $H_1(z)$, en utilisant la méthode de périodogramme, estimer et tracer la DSP à la sortie du filtre.
 - ii) Si le bruit est filtré avec un filtre passe-haut $H_2(z)$, en utilisant la méthode de périodogramme, estimer la DSP et tracer la DSP à la sortie du filtre.
- 6) Un bruit normal avec $m = 0$ et $\sigma^2 = 16$ est filtré en passe-bas puis multiplié par un signal sinusoïdal d'amplitude $A = 1$ et de fréquence $f = 300$ Hz. Estimer et tracer les DSPs à la sortie du filtre et à la sortie du dispositif de produit.
- 7) Étant donné deux signaux normaux indépendants passe-bas et passe-haut $y_1(t)$ et $y_2(t)$, respectivement, les deux avec $m = 0$ et $\sigma_1^2 = 16$, $\sigma_2^2 = 1$. Un signal sinusoïdal $x(t)$ d'amplitude $A = 1$ et de fréquence $f = 250$ Hz est ajouté au signal $y_2(t)$ pour obtenir :

$$y_{22}(t) = x(t) + y_2(t)$$

Estimer et tracer les DSPs des signaux, $y_1(t)$, $y_{22}(t)$ et $y(t)$.

$$y(t) = y_{22}(t) + y_1(t)$$

TP 4

Estimation de la DSP en utilisant les méthodes de Welch, Multitaper, et Burg

But

Le travail pratique 3 a été consacré à l'estimation de la densité spectrale de puissance d'un signal en utilisant un périodogramme simple. Le présent travail pratique porte sur l'estimation de la densité spectrale de puissance d'un signal en utilisant les méthodes de Welch, Multitaper et Burg puis les comparer avec l'estimateur simple.

Fonctions Matlab

rand.m, randn.m, semilogy.m, periodogram.m, pwelch.m, pmtm.m, pburg.

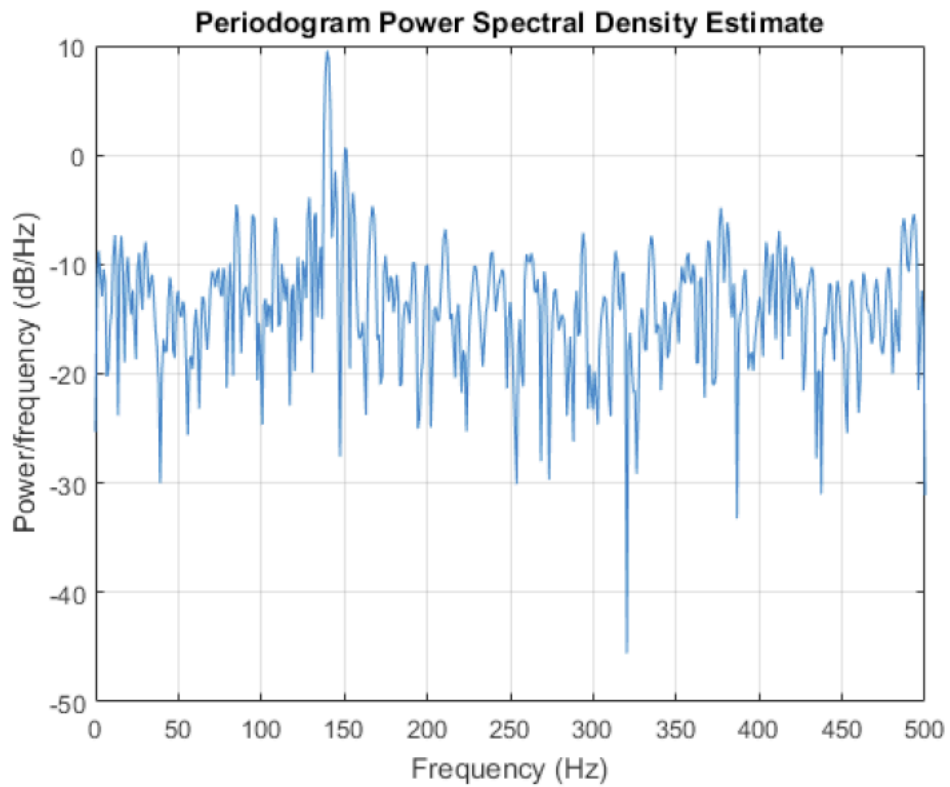
Méthode Welch

Un estimateur amélioré de la DSP est celui proposé par Welch. La méthode consiste à diviser les séries de données du temps en segments, calculant un périodogramme modifié de chaque segment.

Exemple

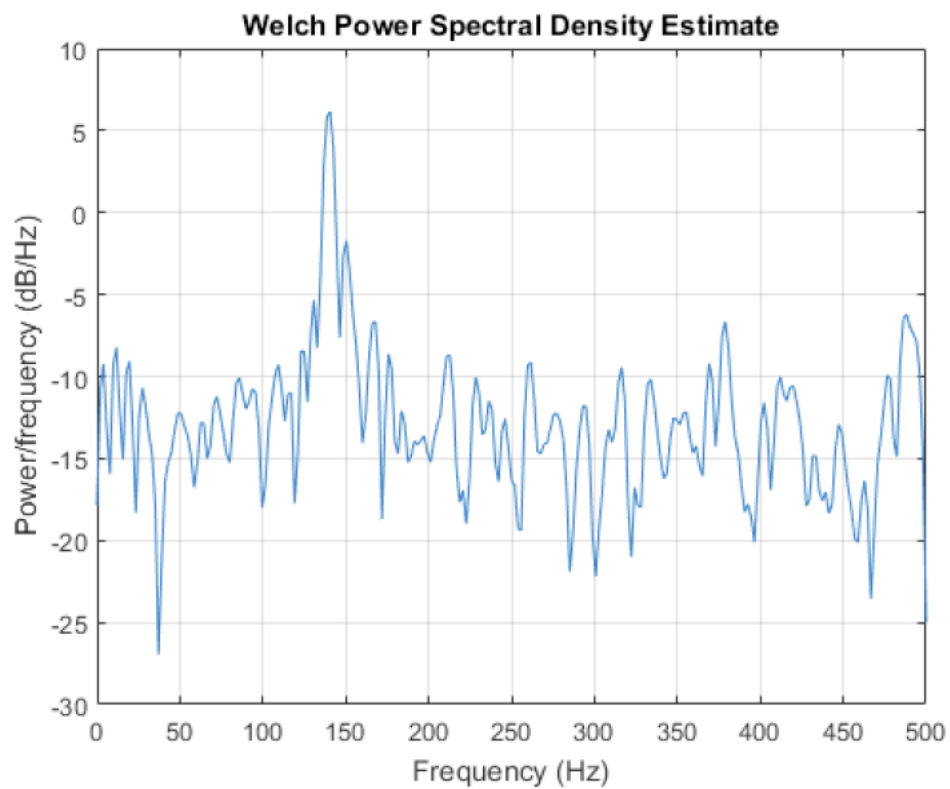
Considérons un signal contenant 301 échantillons.

```
fs = 1000;           % Sampling frequency
t = (0:0.3*fs)/fs;  % 301 samples
A = [2 8];          % Sinusoid amplitudes (row vector)
f = [150;140];      % Sinusoid frequencies (column vector)
xn = A*sin(2*pi*f*t) + 5*randn(size(t));
periodogram(xn,rectwin(length(xn)),1024,fs)
```



Nous pouvons obtenir le spectre estimé de Welch pour 3 segments avec un overlap de 50% en utilisant une fenêtre rectangulaire.

```
pwelch(xn,rectwin(150),50,512,fs)
```



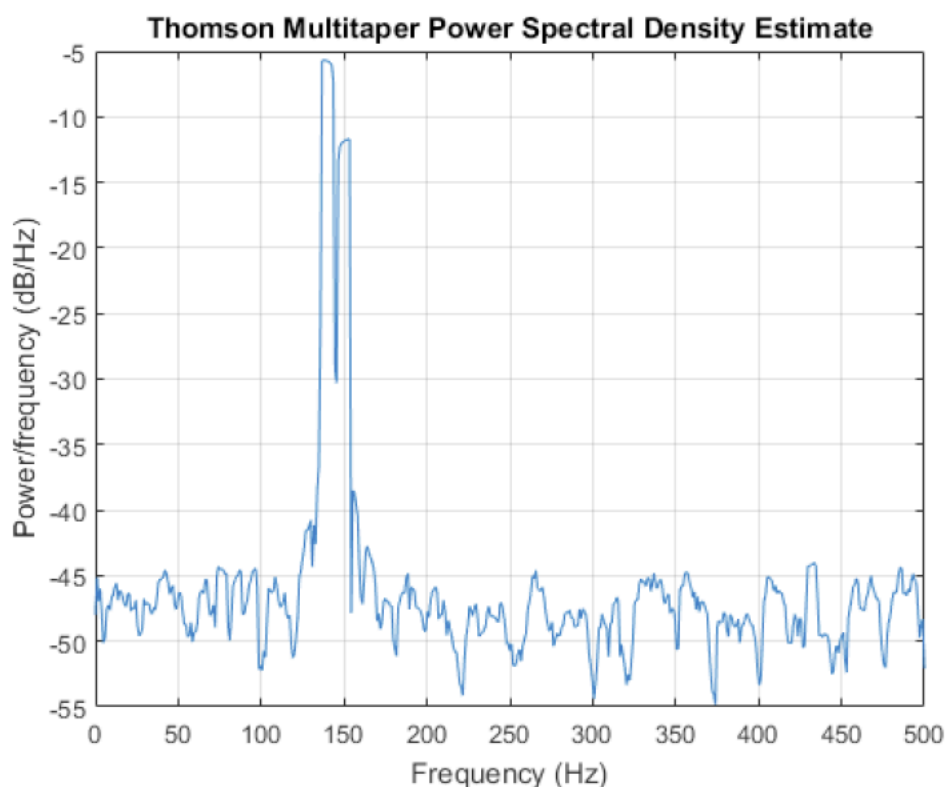
Méthode multitaper

Le périodogramme peut être interprété comme le filtrage d'un signal de longueur L $x_L(n)$ à travers un banc de filtres (un ensemble de filtres en parallèle) composé de L FIR filtres passe bande.

La méthode `pmtm.m` est utilisée pour calculer la DSP d'un signal.

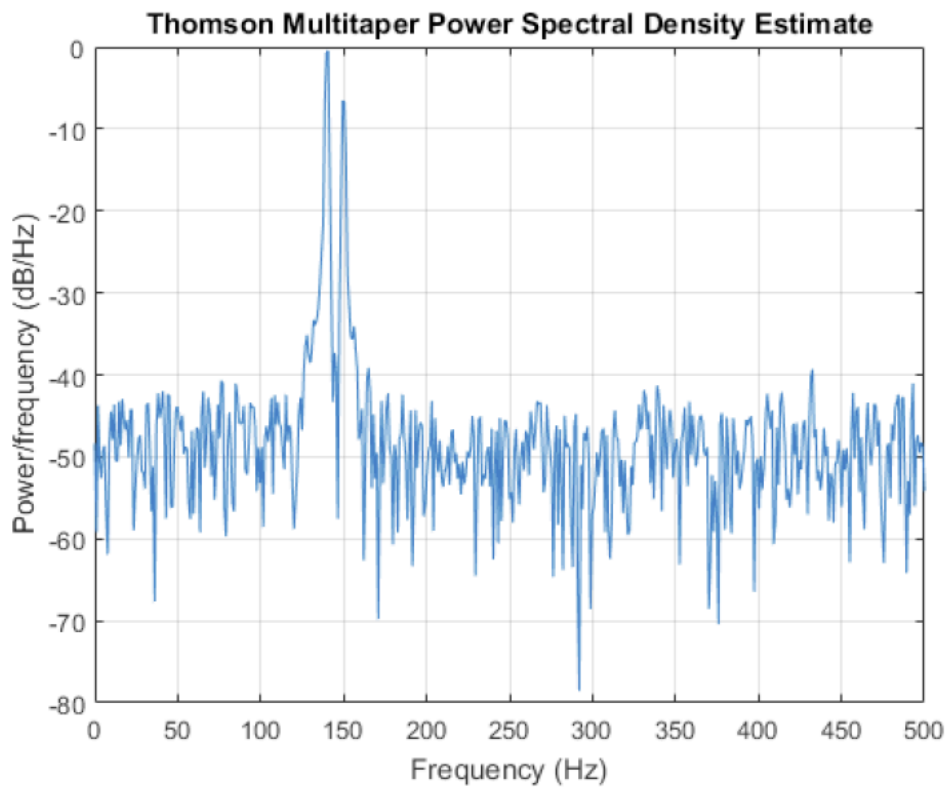
Exemple

```
fs = 1000; % Sampling frequency
t = (0:fs)/fs; % One second worth of samples
A = [1 2]; % Sinusoid amplitudes
f = [150;140]; % Sinusoid frequencies
xn = A*sin(2*pi*f*t) + 0.1*randn(size(t));
pmtm(xn,4,[],fs)
```



En réduisant le produit temps bande passante, vous pouvez augmenter la résolution aux dépens d'une variance plus importante.

```
pmtm(xn,1.5,[],fs)
```



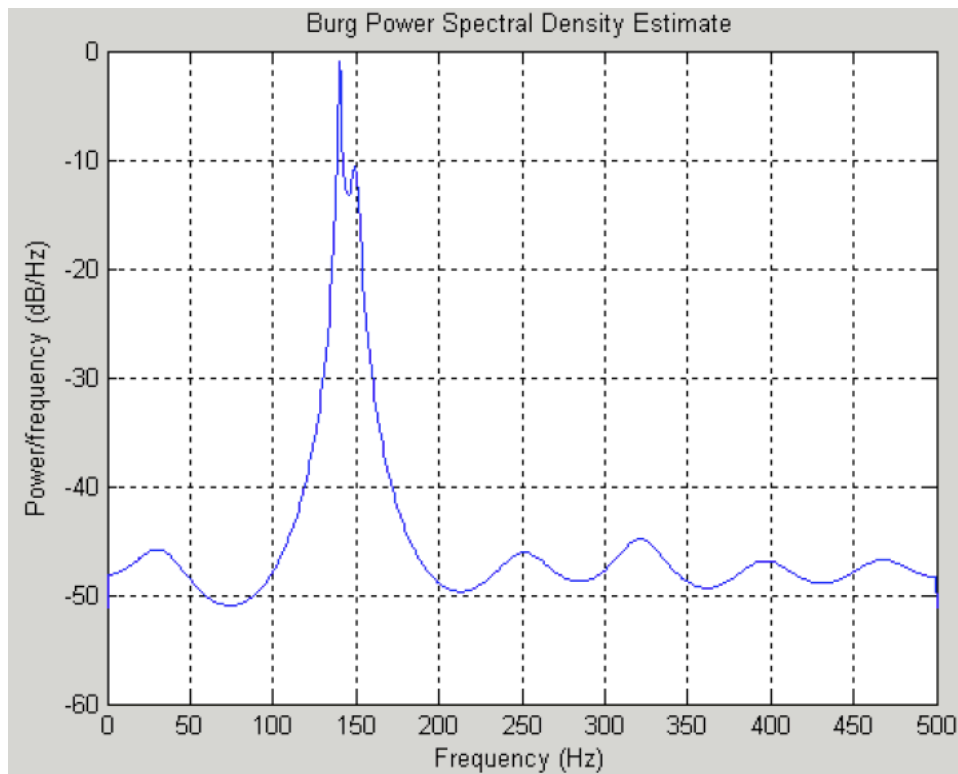
Méthode de Burg

La méthode de Burg pour l'estimation spectrale est basée sur la minimisation des erreurs de la prédiction en boucle ouverte et en boucle fermée en satisfaisant la récursion de Levinson-Durbin. La méthode de Burg évite le calcul de la fonction d'autocorrélation et estime directement les coefficients de réflexion.

Exemple

Calculons l'estimation de la DSP d'un signal bruité en utilisant la méthode de Burg.

```
fs = 1000; % Sampling frequency
t = (0:fs)/fs; % One second worth of samples
A = [1 2]; % Sinusoid amplitudes
f = [150;140]; % Sinusoid frequencies
xn = A*sin(2*pi*f*t) + 0.1*randn(size(t));
pburg(xn,14,1024,fs)
```



Travail demandé

- 1) En utilisant les méthodes de « Welch », « Multitaper » et « Burg », estimer les DSPs et montrer les tracés DSP (utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy) pour deux signaux blanc X1 (normal) et X2 (uniforme), les deux avec $m = 0$ et $\sigma^2 = 1$. Comparer le graphe obtenu avec celui d'un périodogramme simple.
- 2) i) Générer deux signaux X1 (normal) et X2 (uniforme) ayant $m = 0$ et $\sigma^2 = 1$. Tracer les fonctions d'autocorrélation.
 ii) Les signaux sinusoïdaux d'amplitude A et de fréquences $f_1=150\text{Hz}$ et $f_2=250\text{Hz}$ sont ajoutés respectivement aux signaux aléatoires X₁ et X₂ précédents. En utilisant les méthodes de Welch, Multitaper et Burg, estimer leurs DSPs et montrer les tracés DSP pour A = 1 et A = 2. Comparer les tracés avec la méthode du périodogramme simple. (Utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy).
- 3) En utilisant les méthodes de Welch, Multitaper et Burg, estimer les DSPs et montrer les tracés DSP pour les signaux blancs normaux avec :
 1. $m = 4$, et $\sigma^2 = 9$.
 2. $m = 5$ et $\sigma^2 = 16$.

Utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy.

- 4) En utilisant la méthode de Welch, Multitaper et Burg, estimer la DSP et montrer les tracés DSP pour des signaux sinusoïdaux avec $A_1 = 3$, $f_1 = 270$ Hz ainsi que $A_2 = 8$, $f_2 = 175$ Hz dans un bruit blanc normal avec $m = 3$ et $\sigma^2 = 25$, comme représenté sur la Fig. 3.

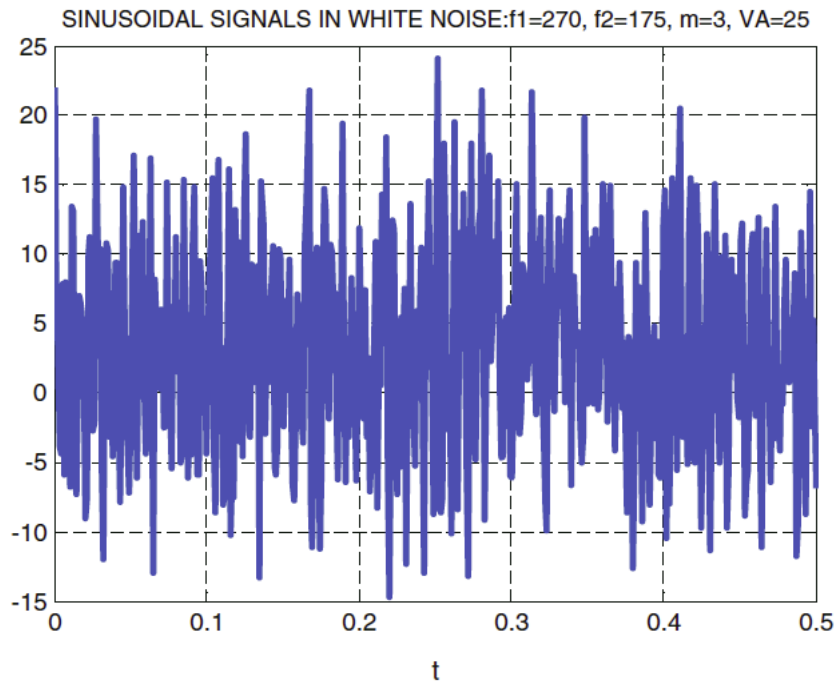


Fig. 3 Signaux sinusoïdaux en bruit blanc à moyenne non nulle

Utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy.

- 5) Étant donné un bruit normal avec $m = 4$ et $\sigma^2 = 4$.
- i) Si le bruit est filtré avec un filtre passe-bas $H_1(z)$, en utilisant la méthode de Welch, Multitaper et Burg, estimer et tracer la DSP à la sortie du filtre.
 - ii) Si le bruit est filtré avec un filtre passe-haut $H_2(z)$, en utilisant la méthode de Welch, Multitaper et Burg, estimer la DSP et tracer la DSP à la sortie du filtre.

Utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy.

- 6) Un bruit normal avec $m = 0$ et $\sigma^2 = 16$ est filtré avec un filtre passe bas puis multiplié par un signal sinusoïdal d'amplitude $A = 1$ et $f = 300$ Hz.

Estimer et tracer les DSPs à la sortie du filtre et à la sortie du dispositif de produit en utilisant la méthode de Welch, Multitaper et Burg. (Utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy).

- 7) Étant donné deux signaux normaux indépendants passe-bas et passe-haut $y_1(t)$ et $y_2(t)$, respectivement, les deux avec $m = 0$ et $\sigma_1^2 = 16$, $\sigma_2^2 = 1$. Un signal sinusoïdal $x(t)$ d'amplitude $A = 1$ et $f = 250$ Hz est ajouté au signal $y_2(t)$ pour obtenir

$$y_{22}(t) = x(t) + y_2(t)$$

Estimer et tracer les DSPs des signaux, $y_1(t)$, $y_{22}(t)$ et $y(t)$ en utilisant la méthode de Welch, Multitaper et Burg (utiliser une échelle logarithmique en ordonnée avec l'instruction semilogy).

$$y(t) = y_{22}(t) + y_1(t)$$