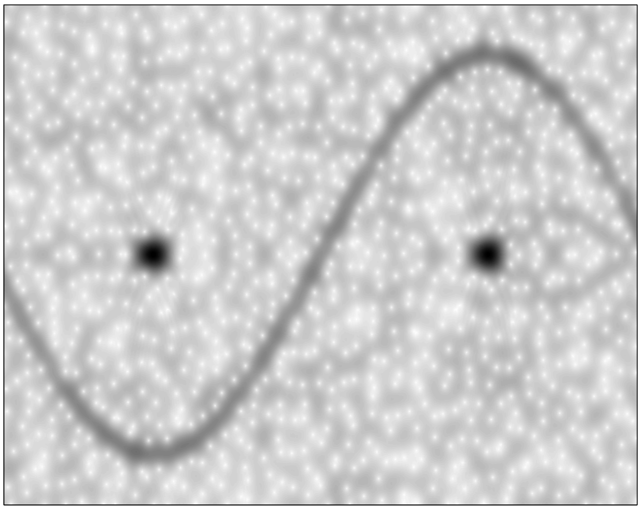
Polycopié du cours de traitement du signal

Nom :

Prénom :

Promo : 2016

Année de l'enseignement : 2020



Ce polycopié, rédigé en avril 2020, est destiné aux élèves de la promo 2016 de l'école Centrale de Pékin. Merci de ne pas diffuser ce document sur le réseau internet. Il s'agit d'un support pour l'enseignement de traitement du signal d'Antoine Roueff, enseignant-chercheur à l'école Centrale de Marseille. L'objectif de cet enseignement est de présenter les éléments clefs du traitement du signal et d'illustrer la méthodologie en science des données. Plusieurs questions sont posées (en rouge) afin de rendre, dans la mesure du possible, la séance de cours à distance **interactive**. De plus, des progammes `Matlab` sont fournis pour illustrer certains aspects pendant le cours ainsi que pour pouvoir réaliser les séances de travaux pratiques.

Programme

0 Introduction

1 Principales représentations en traitement du signal

1.1 Représentation des systèmes de convolution
1.2 Séance de TD sur les systèmes de convolution
1.3 Classification des signaux et numérisation de l'information
1.4 Séance de TD sur l'échantillonnage
1.5 Caractérisation et représentation des fonctions aléatoires
1.6 Séance de TP sur l'analyse spectrale
1.7 Filtrage des fonctions aléatoires
1.8 Séance de TD sur le filtrage des fonctions aléatoires

2 Illustration en science des données : la discrimination

2.1 Discrimination linéaire
2.2 Séance de TP sur la discrimination linéaire
2.3 Méthode probabiliste
2.4 Séance de TP sur la méthode probabiliste
2.5 Analyse de différentes approches possibles sur un exemple
2.6 Séance de TP sur la reconnaissance de chiffres
2.7 Risque de Bayes
2.8 Séance de TP sur le risque de Bayes

0 Introduction

Ce chapitre introductif a pour but de présenter le contexte scientifique ainsi que les objectifs de cet enseignement.

Qu'est-ce que le traitement du signal ?

Qu'est-ce qu'un algorithme ?

En termes mathématiques, on peut représenter une donnée par un vecteur $\boldsymbol{y}$, une information par un vecteur $\boldsymbol{\theta}$ et un algorithme par une fonction f. Le problème consiste à trouver une fonction f telle que

$$\boldsymbol{\theta} = f(\boldsymbol{y})$$

Exemple 1. On considère des photos de chiens et de chats. On veut définir une technique capable de reconnaître si une photo représente un chien ou un chat.

Que proposez-vous pour résoudre ce problème ?

Exemple 2. On considère un signal x échantillonné avec le temps d'échantillonnage T_e (voir figure 1).

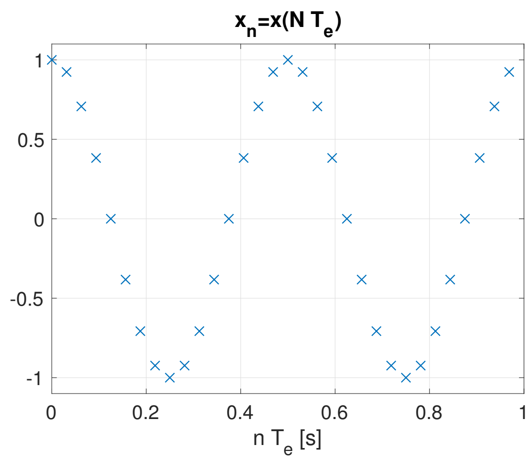


FIGURE 1 – Représentation d'un signal x échantillonné avec un temps d'échantillonnage T_e. Ce graphe est obtenu avec le programme `main_sin.m`.

La donnée d'entrée est $\boldsymbol{y} = (x_n)_n$ et pour t fixé, on cherche l'information $\boldsymbol{\theta} = x(t)$.

Que proposez-vous pour résoudre ce problème ?

Avant de rentrer dans les détails techniques, essayons d'effectuer un petit rappel historique.

Depuis quand le traitement du signal existe ?

A la fin des années 40, la théorie de l'information est découverte. **Quelles en sont les conséquences ?**

A la fin du XXème siècle, le traitement du signal est utilisé dans de nombreux domaines : les communications, le multimédia, la finance, ... **Quelles sont les avancées technologiques qui expliquent cette évolution ?**

Quels sont les changements auxquels nous assistons en ce début de XXIème siècle ?

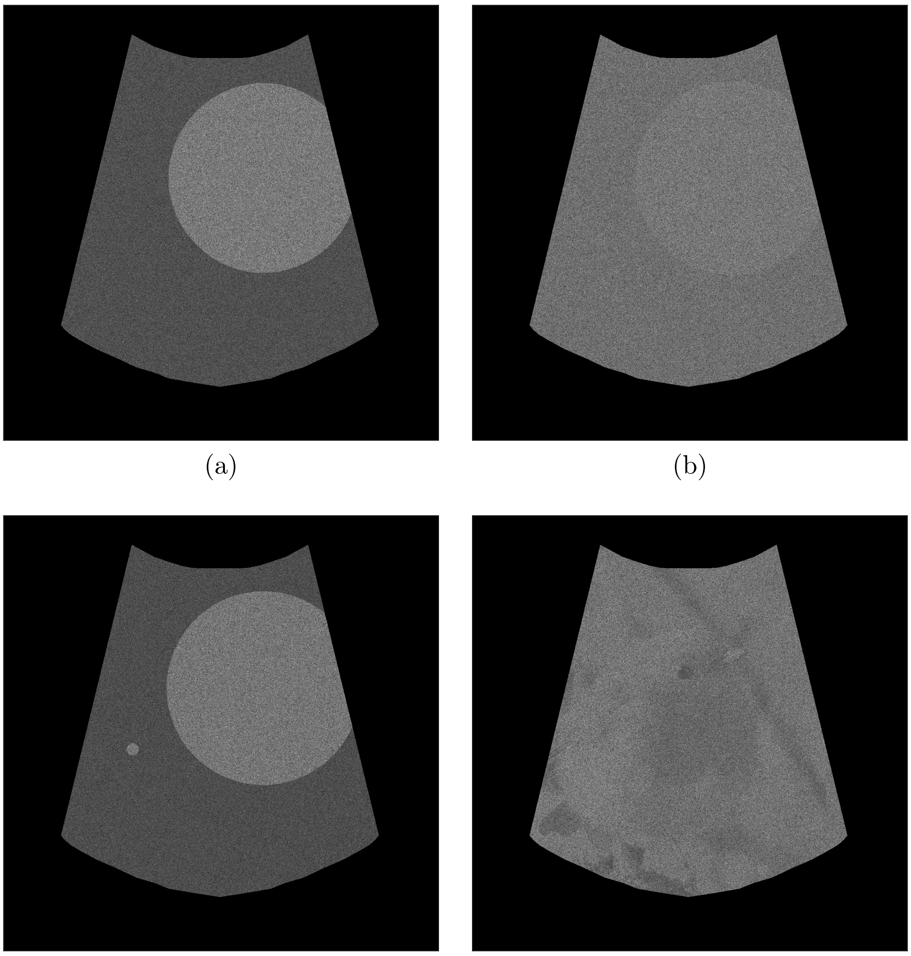
Exemples de domaine d'application ?

Pourquoi est-ce important d'avoir une formation dans ce domaine pour une formation généraliste ?

Quelles sont les compétences requises pour un ingénieur en traitement du signal ?

Illustration des problèmes qui peuvent apparaître en traitement du signal.

Exemple d'un système d'imagerie. Pour simplifier, on suppose que l'organe auquel on s'intéresse ressemble à un disque (voir figure 2(a))



(a) (b) (c) (d)

FIGURE 2 – Exemple d'un système d'imagerie. Lancez le programme `main_speckle` sous `Matlab`.

D'où vient la difficulté de l'analyse de la figure (b) ?

D'où vient la difficulté de l'analyse de la figure (c) ?

D'où vient la difficulté de l'analyse de la figure (d) ?

FIGURE 3 – Le TS consiste à définir un algorithme qui à partir d'une donnée extrait une information.

Le traitement du signal (TS) joue un rôle structurant car pour chaque problème, vous devez choisir la représentation, qui correspond à projeter le point de vue sur le problème considéré, et ceci en fonction des connaissances a priori dont vous disposez sur les mécanismes physiques mis en œuvre et en fonction de la tâche à accomplir. L'objectif de cet enseignement est qu'à l'issue de cet enseignement, les élèves maîtrisent les concepts de bases afin d'être capable d'interagir avec des spécialistes.

1 Principales représentations en traitement du signal

1.1 Représentation des systèmes de convolution

En traitement du signal, l'étude des système (voir figure 4) est fondamentale. Les systèmes sont en particulier utilisés pour décrire les interactions entre le signal mesuré par le capteur et le monde physique.

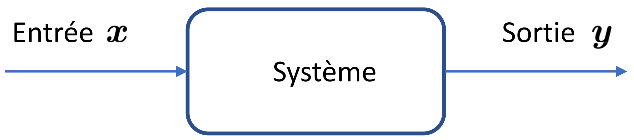


FIGURE 4 – Représentation d'un système avec une entrée $\boldsymbol{x}$ et une sortie $\boldsymbol{y}$.

Exemple ?

Soit $\mathcal{S}$ un **système** avec une entrée x et une sortie y. On utilise la notation $y = \mathcal{S}(x)$ ou $y(t) = \mathcal{S}(x)(t)$.

Le système $\mathcal{S}$ est **linéaire**, si $\forall (a_1,\, a_2) \in \mathbb{C}^2$, $\forall (x_1, x_2)$, $\mathcal{S}(a_1 x_1 + a_2 x_2)(t) = a_1 \mathcal{S}(x_1)(t) + a_2 \mathcal{S}(x_2)(t)$.

Exemple : un système régi par

$$y(t) = a_0 x(t-\tau) + a_1 \frac{dx(t)}{dt} + a_2 x^2(t) + a_3 \log|x(t)|$$

est-il linéaire ?

Le système $\mathcal{S}$ est **stationnaire**, si pour $y_1(t) = \mathcal{S}(x_1)(t)$ et $x_2(t) = x_1(t-\tau)$, on a $y_2(t) = \mathcal{S}(x_2)(t) = y_1(t-\tau)$.

Exemple 1. On considère le système qui consiste à sélectionner la partie d'un signal entre $-\frac{T}{2} < t < \frac{T}{2}$

$$y(t) = \mathrm{rect}_T(t) x(t) \quad \text{où} \quad \mathrm{rect}_T(t) = \begin{cases} 1 & \text{si } -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{sinon} \end{cases}$$

Ce système est-il stationnaire ?

$$y_1(t) = \mathrm{rect}_T(t) x_1(t)$$

$$y_2(t) = \mathrm{rect}_T(t) x_2(t)$$

$$y_2(t) = \mathrm{rect}_T(t) x_1(t-\tau)$$

$$y_1(t-\tau) = \mathrm{rect}_T(t-\tau) x_1(t-\tau)$$

Exemple 2. On considère le système qui consiste à calculer la moyenne glissante d'un signal sur un temps T

$$y(t) = \int_{t-T/2}^{t+T/2} x(u) du$$

Ce système est-il stationnaire ?

$$y_1(t) = \int_{t-T/2}^{t+T/2} x_1(u) du$$

$$y_2(t) = \int_{t-T/2}^{t+T/2} x_2(u) du$$

$$y_2(t) = \int_{t-T/2}^{t+T/2} x_1(u-\tau) du$$

$$y_2(t) = \int_{t-T/2-\tau}^{t+T/2-\tau} x_1(u) du = y_1(t-\tau)$$

On peut montrer que si un système est linéaire, stationnaire et continu alors il existe une fonction h telle que la sortie du système y est égale au **produit de convolution** $\star$ entre l'entrée x et la fonction h

$$y(t) = (x \star h)(t) = \int_{\mathbb{R}} h(t-u)x(u)du \tag{1}$$

Un tel système s'appelle un **système de convolution**. Il est caractérisé par h que l'on appelle la **réponse impulsionnelle du système**.

Pourquoi appeler h la réponse impulsionnelle ?

Soit $\mathcal{D}$ l'espace vectoriel des fonctions indéfiniment différentiables à support compact. On appelle **distribution** sur $\mathbb{R}$, une forme linéaire continue sur $\mathcal{D}$. Soit $a \in \mathbb{R}$, la distribution de **Dirac** δ_a est définie par

$$\forall \phi \in \mathcal{D} \quad \delta_a(\phi) = \phi(a) \tag{2}$$

En traitement du signal, la distribution de Dirac est souvent notée comme une fonction $\delta(t-a)$. Cette notation est utilisée pour mener intuitivement des calculs, qui en toute rigueur devraient être démontrés au sens des distributions. Dans cet esprit, on admettra (et on apprendra par cœur) les relations suivantes, $\forall a \in \mathbb{R}$ et $\forall x$

$$\int_{\mathbb{R}} \delta(t-a)dt = 1, \quad x(t) \,.\, \delta(t-a) = x(a)\,\delta(t-a) \quad \text{et} \quad x(t) \star \delta(t-a) = x(t-a) \tag{3}$$

Si on souhaite effectuer les calculs de manière rigoureuse sans avoir à manipuler les distributions, une solution est d'utiliser la propriété suivante (non démontrée ici)

$$\delta(t) = \lim_{T \mapsto 0} \frac{1}{T} \mathrm{rect}_T(t) \quad \text{où} \quad \mathrm{rect}_T(t) = \begin{cases} 1 & \text{si } -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{sinon} \end{cases}$$

Quelle peut être l'utilité de la distribution de Dirac en traitement du signal ?

Exemple de la prise en photo avec un flou horizontal
On suppose qu'au moment de prendre une photo, on a bougé horizontalement. Chaque ligne de l'image correspondante est un signal à une dimension qui peut être approximée par la sommes ci-dessous

$$y(t) = \int_{t-T/2}^{t+T/2} x(u)du$$

où x est l'image nette, y est l'image floue et T est une constante.

Ce système est-il un système de convolution ?

Pourquoi les systèmes de convolution jouent un rôle important ?

Propriété : les fonctions $t \mapsto e^{2i\pi\nu t}$ sont les fonctions propres des systèmes de convolution.

En effet, si $x(t) = \exp(2i\pi\nu t)$, alors

$$y(t) = \int_{\mathbb{R}} h(u) \exp(2i\pi\nu(t-u))du$$

et

$$y(t) = \exp(2i\pi\nu t) \int_{\mathbb{R}} h(u) \exp(-2i\pi\nu u)du = x(t)H(\nu)$$

où

$$H(\nu) = \int_{\mathbb{R}} h(u)e^{-2i\pi\nu u}du$$

est la valeur propre associée.

Connaissez-vous la transformée qui passe de $h(t)$ à $H(\nu)$?

Quand elle existe, **la transformée de Fourier** d'un signal x est donnée par l'intégrale

$$X(\nu) = \int_{\mathbb{R}} x(t)e^{-2i\pi\nu t}dt \quad \text{où } \nu \text{ est la variable fréquence.} \tag{4}$$

On peut montrer que si $y = x \star h$, alors
$$Y(\nu) = X(\nu)\, H(\nu) \tag{5}$$
où $X(\nu)$, $H(\nu)$ et $Y(\nu)$ sont les transformées de Fourier (que l'on suppose exister) de $x(t)$, $h(t)$ et $y(t)$.

La fonction $H(\nu)$ s'appelle la **fonction de transfert** du système.

Comment démontrer la relation $Y(\nu) = X(\nu)H(\nu)$?

En cours de Mathématique, il est démontré que si les intégrales sont définies, alors la transformée de Fourier inverse existe
$$x(t) = \int_{\mathbb{R}} X(\nu) e^{j2\pi\nu t} d\nu$$
en utilisant la linéarité de la convolution, on obtient
$$y(t) = (x \star h)(t) = \int_{\mathbb{R}} X(\nu)(e^{j2\pi\nu t} \star h)(t) d\nu$$
Or les fonctions $e^{j2\pi\nu t}$ sont les fonctions propres, donc
$$y(t) = \int_{\mathbb{R}} X(\nu)H(\nu)e^{j2\pi\nu t} d\nu$$
Donc par identification, on a
$$Y(\nu) = X(\nu)H(\nu)$$
Ainsi le fait de décomposer nos signaux sur les fonctions propres des systèmes de convolution, permet d'aboutir à une représentation simple où une convolution devient un simple produit.

Pourquoi les systèmes de convolution jouent un rôle important ?
D'une part, leur domaine d'applicabilité est grand et d'autre part leur représentation est simple

Propriétés de la transformée de Fourier

Transformée de Fourier inverse : $x(t) = \int_{\mathbb{R}} X(\nu)e^{2i\pi\nu t} d\nu$

Cette relation, qui est démontrée en cours de mathématique, est très importante car elle montre que la transformation de Fourier est réversible (à condition que l'intégrale soit définie).

Conjugaison : $TF(x^*(t)) = X^*(-\nu)$, où $*$ représente le complexe conjugué. On en déduit que si un signal est réel (au sens où $x(t) \in \mathbb{R}$), alors le module de la TF est pair et la phase est impaire.

Cette relation est facile à démontrer (à faire en exercice si besoin).

Translation : $TF(x(t-\tau)) = X(\nu)e^{-2i\pi\nu\tau}$,

En effet
$$TF(x(t-\tau)) = \int x(t-\tau)\exp(-2i\pi\nu t) dt$$
avec le changement de variable $u = t - \tau$
$$TF(x(t-\tau)) = \int x(u)\exp(-2i\pi\nu(u+\tau)) du = X(\nu)\exp(-2i\pi\nu\tau)$$

Modulation d'amplitude : $TF(x(t)e^{2i\pi\nu_0 t}) = X(\nu - \nu_0)$

Cette transformation est duale de la précédente. Elle correspond à une translation du spectre.

Exemple. Si
$$x(t) = e^{2i\pi\nu_1 t}$$
alors
$$x(t)e^{2i\pi\nu_0 t} = x(t)e^{2i\pi(\nu_0+\nu_1)t}$$

Quand $\nu_1 > 0$, le spectre de x est déplacé vers les hautes fréquences ou vers les basses fréquences ?

Dérivation : $TF\left(\frac{d^n}{dt^n}x(t)\right) = (2i\pi\nu)^n X(\nu)$,

Pour une démonstration rigoureuse, voir cours de mathématique. On insistera ici sur son utilité pour la résolution de système d'équations différentielles.

Exemple.

$$x(t) = y(t) + a\frac{dy}{dt}$$

devient

$$X(\nu) = Y(\nu) + a2i\pi\nu Y(\nu)$$

et ainsi

$$y(t) = TF^{-1}\left(\frac{X(\nu)}{1 + a2i\pi\nu}\right)$$

Relation de Parseval : $\int_{\mathbb{R}} x(t)y^*(t)dt = \int_{\mathbb{R}} X(\nu)Y^*(\nu)d\nu$

Cette relation permet très souvent de simplifier les calculs.

Que reconnaissez-vous dans cette formule ?

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

Pour démontrer cette relation, il faut effectuer un changement de variable.

$$TF(x(at))(\nu) = \int_{\mathbb{R}} x(at)\exp\left(-2i\pi\nu t\right)dt$$

et avec $t' = at$, on a si $a > 0$

$$TF(x(at))(\nu) = \int_{\mathbb{R}} x(t')\exp\left(-2i\pi\nu\frac{t'}{a}\right)dt'/a = \frac{1}{a}X(\nu/a)$$

et si $a < 0$

$$TF(x(at))(\nu) = -\int_{\mathbb{R}} x(t')\left(-2i\pi\nu\frac{t'}{a}\right)dt'/a = \frac{1}{-a}X(\nu/a)$$

On définit les signaux sinus cardinal, rectangle, échelon d'Heaviside et peigne de Diracs :

$$\mathrm{sinc}(t) = \frac{\sin(t)}{t}, \quad \mathrm{rect}_T(t) = \left\{\begin{array}{ll} 1 & \text{si } -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{sinon} \end{array}\right., \quad \theta(t) = \left\{\begin{array}{ll} 1 & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{array}\right., \quad ⨿_T(t) = \sum_{n\in\mathbb{Z}} \delta(t - nT) \tag{6}$$

Exemple de calcul pour lequel la relation de Parseval est utile.

Sachant que $TF(\mathrm{rect}_T(t)) = T\mathrm{sinc}(\pi\nu T)$, montrez que

$$\int_{\mathbb{R}} \mathrm{sinc}^2(\pi\nu T)d\nu = \frac{1}{T}$$

Exemple de calcul de convolution entre x et h pour

$$x(t) = \cos(2\pi\nu_0 t) \quad \text{et} \quad h(t) = e^{-\pi t^2}$$

Montrez que $x(t) = e^{-\pi\nu_0^2}\cos(2\pi\nu_0 t)$.

Un système $\mathcal{S}$ est **causal** si $\forall x$ tel que $x(t) = 0$ pour $t < 0$, on a aussi $y(t) = \mathcal{S}(x)(t) = 0$ pour $t < 0$.

Ce sous-ensemble est très important pour les physiciens car la causalité est une propriété vérifiée par l'ensemble des systèmes physiques. En effet, cela revient à dire que tout système physique ne peut se mettre en mouvement avant d'avoir été excité.

$x(t)$	$X(\nu)$
$\delta(t)$	1
1	$\delta(\nu)$
$\delta(t-\tau)$	$e^{-2i\pi\nu\tau}$
$e^{2i\pi\nu_0 t}$	$\delta(\nu-\nu_0)$
$\cos(2\pi\nu_0 t)$	$\frac{1}{2}\left(\delta(\nu-\nu_0)+\delta(\nu+\nu_0)\right)$
$\sin(2\pi\nu_0 t)$	$\frac{1}{2i}\left(\delta(\nu-\nu_0)-\delta(\nu+\nu_0)\right)$
$\mathrm{rect}_T(t)$	$T\mathrm{sinc}(\pi\nu T)$
$\sqcup\!\sqcup_T(t)$	$\frac{1}{T}\sqcup\!\sqcup_{\frac{1}{T}}(\nu)$
$e^{-\pi t^2}$	$e^{-\pi\nu^2}$
$e^{-\lvert t\rvert}$	$\frac{2}{1+(2\pi\nu)^2}$

Table 1 – Transformées de Fourier usuelles.

On peut montrer qu'un système de convolution est causal, si et seulement si $h(t)=0$ pour $t<0$.

Une autre conséquence de la causalité est décrite par les relations de Kramers-Kronig [1]. Soit H_r et H_i tels que $H(\nu)=H_r(\nu)+iH_i(\nu)$, on a $H_r(\nu)=\frac{1}{\pi}\mathrm{Vp}\left(\int_{\mathbb{R}}\frac{H_i(\xi)}{\xi-\nu}d\xi\right)$ et $H_i(\nu)=-\frac{1}{\pi}\mathrm{Vp}\left(\int_{\mathbb{R}}\frac{H_r(\xi)}{\xi-\nu}d\xi\right)$, où $\mathrm{Vp}\int_{\mathbb{R}}.dt=\lim_{\epsilon\mapsto 0}\left(\int_{-\infty}^{-\epsilon}.dt+\int_{\epsilon}^{\infty}.dt\right)$. Ces intégrales sont à considérer au sens des distributions.