Science des Données

Gabriel Peyré

10 septembre 2024

Table des matières

1 Théorie de Shannon

1.1 Signaux Analogiques et Discrets

Définition 1.1 Un signal analogique est une fonction $f_0 \in \mathcal{L}^2([0,1])$ où $[0,1]$ est le domaine d'acquisition. Une image analogique est une fonction $2D$ $f_0 \in \mathcal{L}^2\left([0,1]^2\right)$.

La plupart des méthodes discutées dans ce cours s'étendent à des fonctions multi-dimensionnelles

$$f_0 : [0,1]^d \longrightarrow [0,1]^s$$

où d est la dimension de l'espace d'entrée et s est la dimension de l'espace de sortie.

1.1.1 Acquisition et Échantillonnage

Définition 1.2 L'acquisition du signal est une projection en petite dimension d'un signal continu effectuée par un outil matériel. L'opération d'échantillonage correspond donc à une correspondance de l'ensemble des fonctions continues à un vecteur discret de dimension fini :

$$f_0 \in \mathcal{L}^2\left([0,1]^d\right) \longmapsto f \in \mathbb{C}^N$$

1.1.2 Echantillonneur Invariant par Translation

Définition 1.3 Un échantillonneur invariant par translation performe l'acquisition comme la convolution entre le signal continu et une impulsion constante de réponse h translaté à l'emplacement d'échantillonnage :

$$f_n = \int_{-S/2}^{S/2} f_0(x) h\left(\frac{n}{N} - x\right) \mathrm{d}x = f_0 \star h\left(\frac{n}{N}\right) \tag{1}$$

1.2 Théorème d'Échantillonage de Shannon

1.2.1 Rappels sur la Transformée de Fourier

Définition 1.4 Pour $f \in \mathcal{L}^1(\mathbb{R})$, on définit sa transformée de Fourier comme :

$$\forall \omega \in \mathbb{R}, \hat{f}(\omega) = \int_{\mathbb{R}} f(x) e^{-ix\omega} \mathrm{d}x \tag{2}$$

Proposition 1.1 On a :

$$\left\|\hat{f}\right\|^2 = (2\pi)^{-1} \|f\|^2$$

de telle sorte que $f \mapsto \hat{f}$ peut être étendu par continuité à tout $\mathcal{L}^2(\mathbb{R})$ ce qui correspond à calculer $\hat{f}$ comme la limite quand $T \to +\infty$ de $\int_{-T}^{T} f(x) e^{-ix\omega} \mathrm{d}x$. Quand de plus $\hat{f} \in \mathcal{L}^1(\mathbb{R})$ on peut inverser la transformée de Fourier :

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) e^{ix\omega} \mathrm{d}x$$

Proposition 1.2 La transformée de Fourier $\mathcal{F} : f \mapsto \hat{f}$ échange la régularité avec la décroissance. Si $f \in \mathcal{C}^p(\mathbb{R})$ avec une transformée de Fourier intégrable, alors $\mathcal{F}\left(f^{(p)}\right)(\omega) = (i\omega)^p \hat{f}(\omega)$ de telle sorte que $\left|\hat{f}(\omega)\right| = \mathcal{O}\left(|\omega|^{-p}\right)$. Réciproquement,

$$\int_{\mathbb{R}} (1 + |\omega|)^p \left|f(\hat{\omega})\right| \mathrm{d}\omega < +\infty \Longrightarrow f \in \mathcal{C}^p(\mathbb{R})$$

1.2.2 Rappels sur la Série de Fourier

On note $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ le tore.

Définition 1.5 Une fonction $f \in \mathcal{L}^2(\mathbb{T})$ est 2π-périodique et peut être vue comme une fonction f de $\mathcal{L}^2([0, 2\pi])$

dont les coefficients de Fourier sont :

$$\forall n \in \mathbb{Z}, \hat{f}_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ixn} \, \mathrm{d}x$$

Cette formule est équivalente au produit scalaire $\hat{f}_n = \langle f, e_n \rangle$ pour le produit scalaire canonique sur les fonctions sur le tore. Pour ce produit scalaire, les $(e_n)_{n\in\mathbb{Z}}$ forment une base de Hilbert. On a donc :

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n \rangle \, e_n$$

et donc :

$$\left\| f - \sum_{n=-N}^{N} \langle f, e_n \rangle \, e_n \right\|_{L^2(\mathbb{T})} \xrightarrow[N\to+\infty]{} 0$$

Proposition 1.3 Cette formule est équivalente au produit scalaire $\hat{f}_n = \langle f, e_n \rangle$ pour le produit scalaire canonique sur les fonctions sur le tore. Pour ce produit scalaire, les $(e_n)_{n\in\mathbb{Z}}$ forment une base de Hilbert. On a donc :

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n \rangle \, e_n$$

et donc :

$$\left\| f - \sum_{n=-N}^{N} \langle f, e_n \rangle \, e_n \right\|_{L^2(\mathbb{T})} \xrightarrow[N\to+\infty]{} 0$$

1.2.3 Formule de Poisson

La formule de Poisson connecte la transformée de Fourier et les séries de Fourier aux opérateurs d'échantillonnage et de périodisation. Pour une fonction $h(\omega)$ (généralement $h = \hat{f}$), sa périodisation est :

$$h_P(\omega) = \sum_n h(\omega - 2\pi n)$$

Cette formule fait sens si $h \in \mathcal{L}^1(\mathbb{R})$ auquel cas $\|h_P\|_{\mathcal{L}^1(\mathbb{T})} \leq \|h\|_{\mathcal{L}^1(\mathbb{R})}$. La formule de Poisson correspond au diagramme suivant :

$$\begin{array}{ccccc} & f(x) & \xrightarrow{\qquad\mathcal{F}\qquad} & \hat{f} & \\ \text{Échantillonnage} & \downarrow & & \downarrow & \text{Périodisation} \\ & (f(n))_n & \xrightarrow{\text{Série de Fourier}} & \sum_n f(n) e^{-i\omega n} & \end{array}$$

Proposition 1.4 — Formule de Poisson. Si $\hat{f}$ est à support compact et $|f(x)| \leq C(1 + |x|)^{-3}$ pour une constante C, alors :

$$\forall \omega \in \mathbb{R}, \sum_n f(n) e^{-i\omega n} = \hat{f}_P(\omega)$$

Démonstration. Puisque $\hat{f}$ est à support compact, $\hat{f}_P$ est bien défini et f étant rapidement décroissante, $\left(\hat{f}\right)_P$ est $\mathcal{C}^1$. Ainsi, on a :

$$\left(\hat{f}\right)_P(\omega) = \sum_k c_k e^{ik\omega}$$

où

$$c_k = \frac{1}{2\pi}\int_0^{2\pi}\left(\hat{f}\right)_P(\omega)\,e^{-ik\omega}\,\mathrm{d}\omega = \frac{1}{2\pi}\int_{\mathbb{R}}\hat{f}(\omega)\,e^{-ik\omega}\,\mathrm{d}\omega = f(-k)$$

puisque

$$\int_0^{2\pi}\sum_n\left|\hat{f}(\omega-2\pi n)\,e^{-ik\omega}\right|\mathrm{d}\omega = \int_{\mathbb{R}}\left|\hat{f}\right|$$

qui est bornée et où on a utilisé la proposition 1.1 puisque $\hat{f}\in\mathcal{L}^1(\mathbb{R})$. ∎

1.2.4 Théorème de Shannon

Le théorème de d'échantillonnage de Shannon énonce une condition suffisante pour que l'opérateur d'échantillonage $f\mapsto(f(ns))_n$ soit inversible pour un certain pas d'échantillonage $s>0$. Elle requiert que $\operatorname{supp}\left(\hat{f}\right)\subset[-\pi/s,\pi/s]$, ce qui implique, par la formule 1.1.

Théorème 1.1 Si $|f(x)|\leq C(1+|x|)^{-3}$ pour une constante C et $\operatorname{supp}\left(\hat{f}\right)\subset[-\pi/s,\pi/s]$ alors on a :

$$\forall x\in\mathbb{R}, f(x)=\sum_n f(ns)\operatorname{sinc}(x/s-n)$$

et la convergence est uniforme.

Démonstration. Le changement de variable $g=f(s\cdot)$ résulte en $\hat{g}=\frac{1}{s}\hat{f}\left(\frac{\cdot}{s}\right)$:

$$\hat{g}(\omega)=\int f(sx)e^{-i\omega x}\,\mathrm{d}x=\frac{1}{s}\int f(sx)e^{-i\left(\frac{\omega}{s}\right)sx}\,\mathrm{d}(sx)=\frac{\hat{f}\left(\frac{\omega}{s}\right)}{s}$$

donc on peut se restreindre à $s=1$. L'hypothèse de support compact implique $\hat{f}(\omega)=\mathbb{1}_{[-\pi,\pi]}(\omega)\,\hat{f}_P(\omega)$. En combinant la formule d'inversion 1.1 et la formule de Poisson 1.4, on obtient :

$$f(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{f}_P(\omega)\,e^{i\omega x}\,\mathrm{d}\omega=\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_n f(n)e^{i\omega(x-n)}\,\mathrm{d}\omega$$

Puisque f est à décroissance rapide, l'intégrale de droite converge absolument et on peut échanger intégrale et somme et obtenir :

$$f(x)=\sum_n f(n)\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{i\omega(x-n)}\,\mathrm{d}\omega=\sum_n f(n)\operatorname{sinc}(x-n)$$

∎