

Analyse Hilbertienne : Transformée de Fourier et ondelettes

Minh-Tâm Tran, Victor Ripplinger

28/02/2013

Table des matières

1 Les données	1
1.1 Formation d'une image	1
1.2 Des données vectorielles aux données fonctionnelles	1
2 Produit scalaire	2
3 DCT : Transformation en cosinus discret	2
4 Séries de Fourier	3
4.1 Rappels	3
4.2 Bases Hilbertiennes, série de Fourier	3
4.3 Base complexe	3
4.4 Comment calculer une série de Fourier d'une fonction ?	3
4.5 Des données spectrales aux valeurs discrètes	4
5 Analyse par ondelettes	4
5.1 Base intuitive	4
5.2 Analyse multirésolution	4

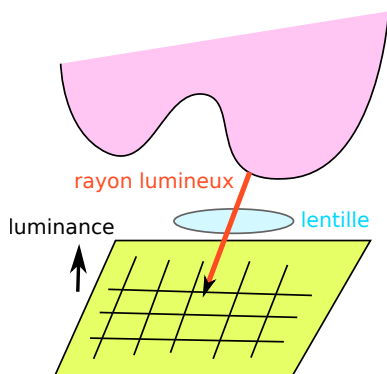
Résumé

Cours avec Sandrine Mouysset, trop bien!!!

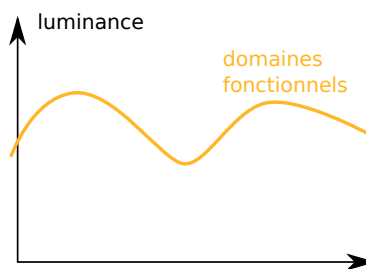
1 Les données

1.1 Formation d'une image

L'ACP est un bon outil mais est clairement un outil de méthode descriptive.



1.2 Des données vectorielles aux données fonctionnelles



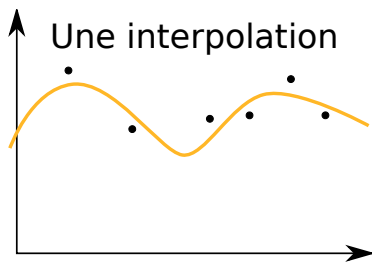
Ici les ensembles de données considérées regroupent des individus qui ne sont pas des vecteur de \mathbb{R}^p mais qui sont des fonctions. Ces fonctions dépendent d'un indice (le temps, l'espace le plus souvent) évoluant dans un intervalle de valeurs $[a, b]$ de \mathbb{R} .

En pratique, les fonctions sont observées en des temps de discrétisation qui peuvent être équi-réparties (ou non), identiques (ou non) pour chaque individu.

Pour rendre les caractère fonctionnel à des données discrétisées, on peut :

- interpoler (Lagrange, Hermite, etc.) ;

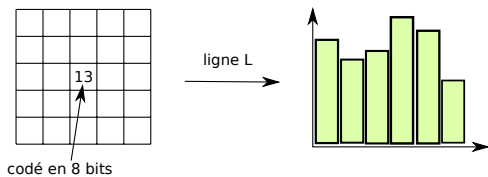
- lisser (B-spline, etc.);
- approcher (Bézier, etc.);
- décomposer (Fourier, ondelettes, etc.).



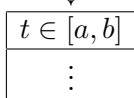
Le prolongement des données discrétisées dans des espaces fonctionnels est particulièrement intéressant dans 2 situations :

- lorsque les fonctions (ou les courbes) sont très régulières;
 - Les méthodes de décomposition permettent la réduction de dimension
- lorsque les données sont bruitées.
 - Un débruitage peut être réalisé par lissage, par seuillage

image numérique luminosité échantillonnée



⋮	⋮
Indicateur X_i	p variables $X_i \in \mathbb{R}^p$
⋮	⋮

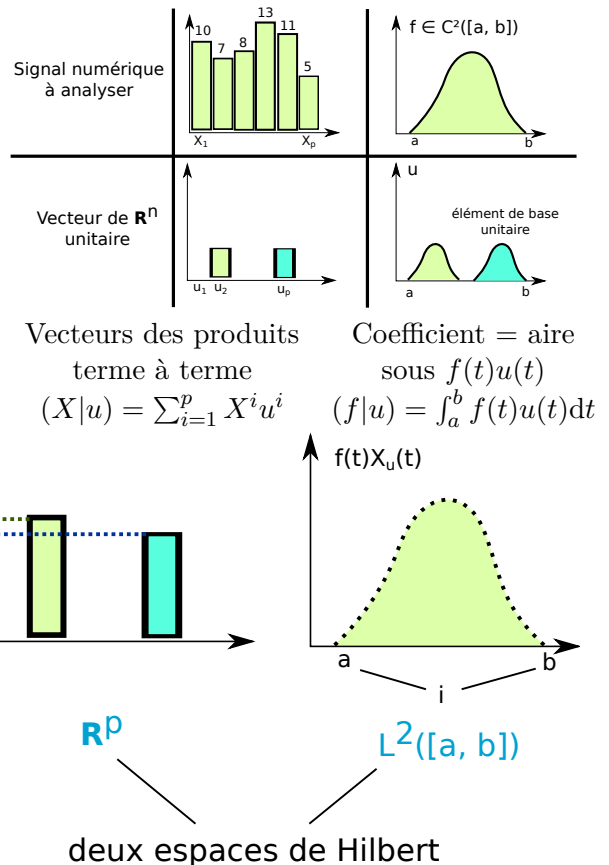


2 Produit scalaire

La difficulté majeure pour rendre leur caractère fonctionnel aux observations discrétisées est le passage

- des espaces de données vectorielles de dimension finie
- aux espaces de données fonctionnelles de dimension infinie

Il faut en particulier généraliser les mécanismes de décomposition (de projection) et redéfinir le produit scalaire.



Dimension finie :

Exemple : \mathbb{R}^p

base orthonormée $u^1 \dots u^p$

$$(u_i | u_i) = 1$$

$$(u_i | u_j) = 0 \text{ si } i \neq j$$

⇔ vecteurs propres de l'ACP

$$f = \sum_{i=1}^p (f | u^i) \cdot u^i$$

Dimension infinie

Exemple : $L^2([a, b])$

Famille dénombrable génératrice de vecteurs orthonormés $u_{i \in \mathbb{N}}$

Base hilbertienne

$$f = \sum_{i \in \mathbb{N}} (f | u_i) u_i$$

$$= \lim_{n \rightarrow \infty} \sum_{i=1}^n (f | u_i) u_i$$

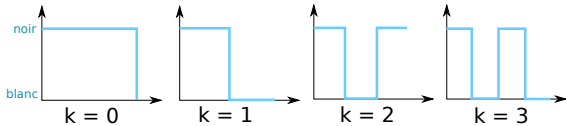
Exemple : Mécanisme d'approximation : on tronque la série

3 DCT : Transformation en cosinus discret

La DCT exprime une suite de nombreux points en termes de somme de fonctions cosinus oscillant à différentes fréquences

$$\text{DCT : } X_k = \sum_{n=0}^{N-1} x_n \cos(\pi/N(n + 1/2)k)$$

→ Utilisée dans les compressions d'image JPEG, MJPEG, MPEG.



Une image est décomposée en vignette de 8×8 pixels. Chaque étape de gauche à droite et de haut en bas est une augmentation de séquence par $1/2$ cycle.

La donnée source (8×8) est transformée en une combinaison linéaire des 64 carrés de fréquence.

Le premier carré correspond au niveau de gris moyen.

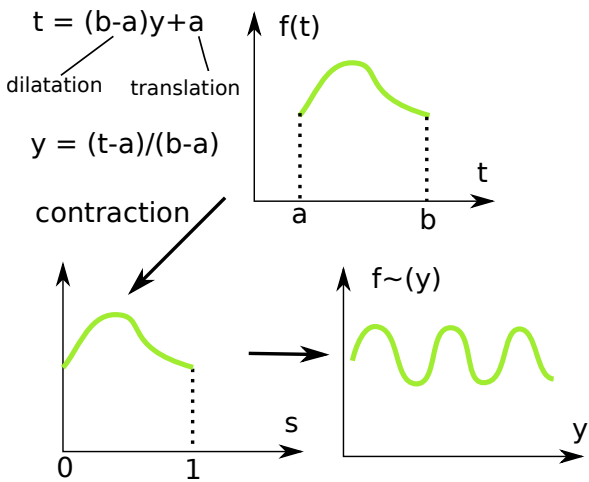
Le calcul des coefficients dans cette base se fait par produit scalaire.

4 Séries de Fourier

4.1 Rappels

Idee : Une fonction périodique $C^1_{p_m}$ peut être décomposée comme une somme de sinusoides (sinus et cosinus) de différentes fréquences.

Préalable : On peut rendre une fonction $f : [a, b] \rightarrow \mathbb{R}$ 1-périodique facilement.



4.2 Bases Hilbertiennes, série de Fourier

La famille $(1, \sqrt{2}\cos(2\pi x), \sqrt{2}\sin(2\pi x), \dots, \sqrt{2}\cos(2\pi n x), \sqrt{2}\sin(2\pi n x), \dots)$ est une base Hilbertienne de $\mathbb{L}^2_p([0, 1])$ On peut donc écrire pour toute fonction $f \in \mathbb{L}^2_p([0, 1])$ dans cette base :

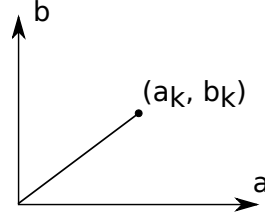
$$f = \sum_{k \in \mathbb{N}} (f | \Phi_k) \Phi_k$$

$$\begin{aligned} \text{d'où } f(t) &= \int_0^1 f(u) \cdot 1 du \\ &+ \sum_{n \in \mathbb{N}^*} [\sqrt{2} \int_0^1 f(u) \cos(2\pi n u) du] \sqrt{2} \cos(2\pi n t) \\ &+ [\sqrt{2} \int_0^1 f(u) \sin(2\pi n u) du] \sqrt{2} \sin(2\pi n t) \\ f(t) &= a_0/2 + \sum_{n \geq 1} [a_n \cos(2\pi n t) + b_n \sin(2\pi n t)] \end{aligned}$$

$$\text{avec } \begin{cases} a_n = 2 \int_0^1 f(u) \cos(2\pi n u) du \\ b_n = 2 \int_0^1 f(u) \sin(2\pi n u) du \end{cases}$$

Les coefficients de la série de Fourier a_n et b_n décomposent / analysent la donnée fonctionnelle uniquement fréquentiellement.

L'information temporelle semble être cachée. Où est l'information sur le temps ? Dans le décalage des sinusoides (c'est à dire les phases).



$$\text{Amplitude : } \sqrt{a_k^2 + b_k^2}$$

$$\text{Phase : } \theta_k \text{ angle, avec } \cos \theta_k = \frac{a_k}{\sqrt{a_k^2 + b_k^2}} \text{ et}$$

$$\sin \theta_k = \frac{b_k}{\sqrt{a_k^2 + b_k^2}}$$

Terme de reconstruction :

$$a_k \cos(2\pi k t) + b_k \sin(2\pi k t) = \sqrt{a_k^2 + b_k^2} \cos(2\pi k t - \theta_k)$$

4.3 Base complexe

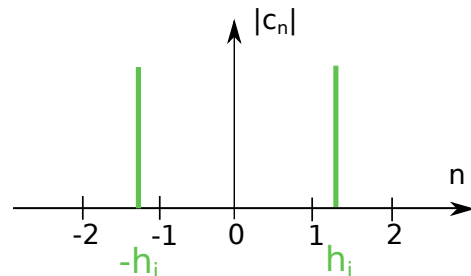
En remplaçant les $\cos(i\pi n t)$ par $\frac{e^{2i\pi n t} + e^{-2i\pi n t}}{2}$ et les $\sin(i\pi n t)$ par $\frac{e^{2i\pi n t} - e^{-2i\pi n t}}{2i}$, on a :

$$\begin{aligned} f(t) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k \cos(2\pi k t) + b_k \sin(2\pi k t)}{\frac{e^{2i\pi k t} + e^{-2i\pi k t}}{2} + \frac{e^{2i\pi k t} - e^{-2i\pi k t}}{2i}} \right] \\ f(t) &= \frac{a_0}{2} + \sum_{n \in \mathbb{Z}} C_n e^{2i\pi n t} \end{aligned}$$

D'où $f(t) = \frac{a_0}{2} + \sum_{n \in \mathbb{Z}} C_n e^{2i\pi n t}$ avec $C_n = \int f(u) e^{-2i\pi n u} du \rightarrow$ fréquence n .

Les c_n sont les coefficients de Fourier complexes.

On peut afficher leur module « spectre d'amplitude » ou leur spectre de phase.



4.4 Comment calculer une série de Fourier d'une fonction ?

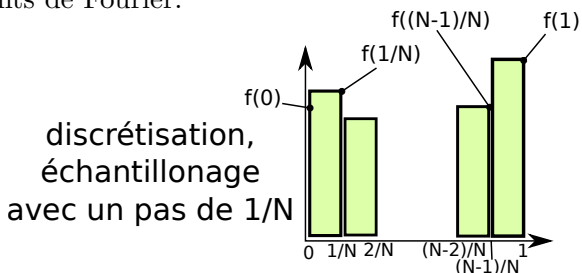
$f(t) = \sum_{n \in \mathbb{Z}} C_n(t) e^{2i\pi n t}$ à partir de la transformée de Fourier de \tilde{f} (plongement nul à support compact).

$$\hat{\tilde{f}}(\lambda) = \int_{\mathbb{R}} \tilde{f}(t) e^{-2i\pi \lambda t} dt$$

$c_n = \hat{f}(n)$ avec n une fréquence entière.

4.5 Des données spectrales aux valeurs discrètes

On va approcher numériquement les coefficients de Fourier.



$$c_n = \int_0^1 f(x) e^{-2i\pi n x} dx \approx \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{j}{N}\right) e^{-2i\pi n \frac{j}{N}} = c_n^N.$$

C'est la TFD (Transformée de Fourier Discrète)

On restreint ainsi le nombre de coefficients. \Rightarrow algorithme FFT (Fast Fourier Transform)

5 Analyse par ondelettes

Somme de Fourier / Transformée de Fourier \rightarrow analyse purement fréquentielle.

Pour rajouter une analyse temporelle, la Transformation de Fourier par fenêtre (glissante) a été définie. Elle permet de localiser simultanément en temps et en fréquence un signal en l'observant sur une fenêtre que l'on translate. \rightarrow transformée de Gabor.

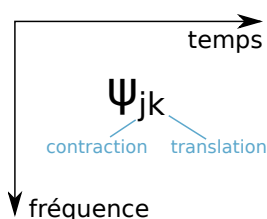
Problème : Tout signal plus court que le support de la fenêtre est sous-localisé dans le temps. \rightarrow obtention d'une réponse localisée dans le temps, implique l'introduction d'une échelle de grandeurs dans l'analyse (taille de la fenêtre). \rightarrow transformation en ondelettes.

Idee : établir un compromis entre temps et fréquence. Si on utilise un signal de base analysant

- assez bien localisé temporellement
- assez bien localisé fréquentiellement
- dont le spectre d'amplitude est bien localisé

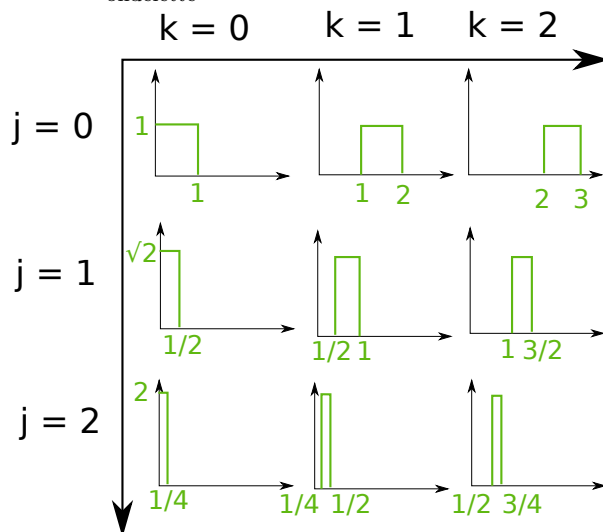
Ondelettes : petites vagues oscillant à un endroit donné approximativement, oscillant à une fréquence donnée approximativement.

5.1 Base intuitive

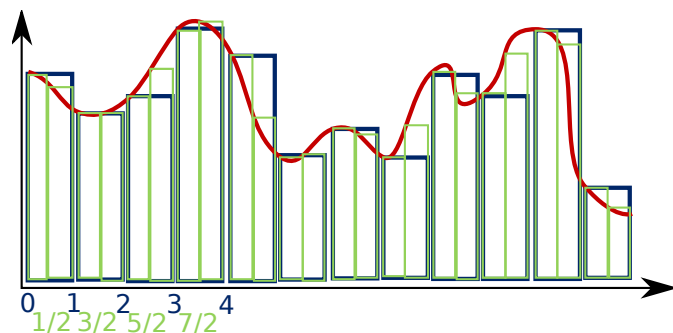


où $f = \sum_j \sum_k (f/\psi_{jk}) \psi_{jk}$ avec j la fréquence et k le temps.

$$\underbrace{2^{j/2}}_{\text{normalisation}} \underbrace{\psi}_{\text{ondelette}} \left(\underbrace{2^j}_{\text{contraction}} x - \underbrace{k}_{\text{translation}} \right).$$



5.2 Analyse multirésolution



Niveau grossier de résolution :

- approximant de niveau 0
- fonction escalier de valeurs $f(-1), f(0), f(1) \dots$

Niveau plus fin de résolution (deux fois plus fin) :

- approximant de niveau 1
- fonction escalier de valeurs : $f(-1), f(-1/2), f(0), f(1/2) \dots$

Soient v_j le sous-espace de fonctions en escalier sur la grille de finesse 2^{-j} . La suite de sous-espace est une Analyse Multirésolution (AMR).

Propriétés :

- Seule la fonction nulle appartient à tous les $v_j : \cap_j v_j = \{0\}$.
- L'approximation à une échelle donnée contient toute l'information contenue aux résolutions plus grossières : $v_0 \subset v_1 \subset v_2 \dots$
- On peut approcher avec une précision arbitraire n'importe quelle fonction : $\lim_{i \rightarrow \infty} v_i = \mathcal{L}^2(\mathbb{R})$.
- On a aussi :

$f \in v_0 \leftrightarrow f(x - k) \in v_0, \forall k \in \mathbb{Z}$ (translatée de k)

$f \in v_i \leftrightarrow f(2x) \in v_{i-1}$ (contractée de 2)

Notons $V_1 = V_0 \oplus W_0$ avec W_0 l'espace de détails qu'il faut rajouter à un approximant de niveau 0 (v_0) pour passer à un approximant de niveau 1. De même $V_{j+1} = V_j \oplus W_j$.

- ψ l'ondelette mère choisie telle que $\{\psi(x - k)\}_{k \in \mathbb{Z}}$ forme une base hilbertienne de W_0 et $\psi_{jk} = 2^{j/2}\psi(2^j x - k)$ forme une base hilbertienne de W_j .
- On sait aussi choisir une fonction ϕ (fonction d'échelle ou ondelette père) dont les translatées forment une base hilbertienne de V_0 .

	Fonctions	Espaces	Bases	j augmente	j diminue
Approximation	fonction échelle ϕ	V_j	$\{\phi_{jk}\}_{jk}$	plus fin	plus grossier
Détail	ondelettes ψ	W_j	$\{\psi_{jk}\}_{jk}$		