

Compression et Débruitage avec les ondelettes.

Anestis Antoniadis

antonia@imag.fr

<http://www-lmc.imag.fr/SMS>

Plan

Généralités sur les décompositions en ondelettes

- Les ondelettes
 - ▷ Analyses multirésolution
 - ▷ Algorithmes de calcul
- Approximation et compression
 - ▷ Linéaire
 - ▷ Non linéaire

Débruitage et problèmes d'estimation non paramétrique

- Régularisation
- Propriétés asymptotiques
- Applications

Généralités

Tant en mathématique qu'en physique, on a souvent intérêt à représenter un "signal" observé comme une superposition adéquate de fonctions élémentaires, cette représentation étant adaptée au problème étudié.

Afin d'obtenir de telles représentations qui ne soient pas uniquement intéressantes pour la théorie mais aussi pour les applications, on a besoin d'algorithmes rapides de calcul.

Une fois en possession de ces "fonctions élémentaires", on peut essayer d'obtenir une bonne reconstitution de l'objet étudié en n'utilisant qu'un nombre limité de telles fonctions, ce qui peut se concevoir comme une tâche d'approximation ou de compression.

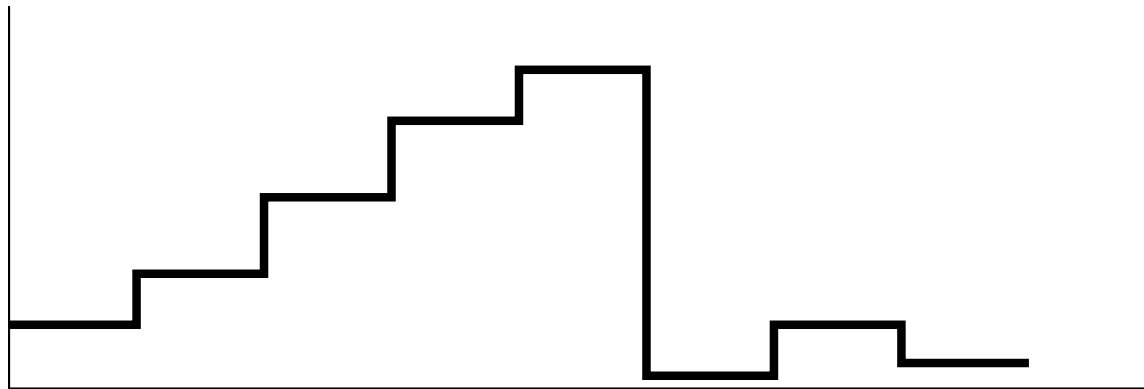
La base de Haar

La base de Haar est la base d'ondelettes la plus simple.

Nous allons présenter dans un premier temps une méthode naturelle de décomposition d'une fonction intégrable.

Imaginons que l'on se donne une fonction $f \in L^1([0,1])$, digitalisée (discrétisée) sur 8 valeurs :

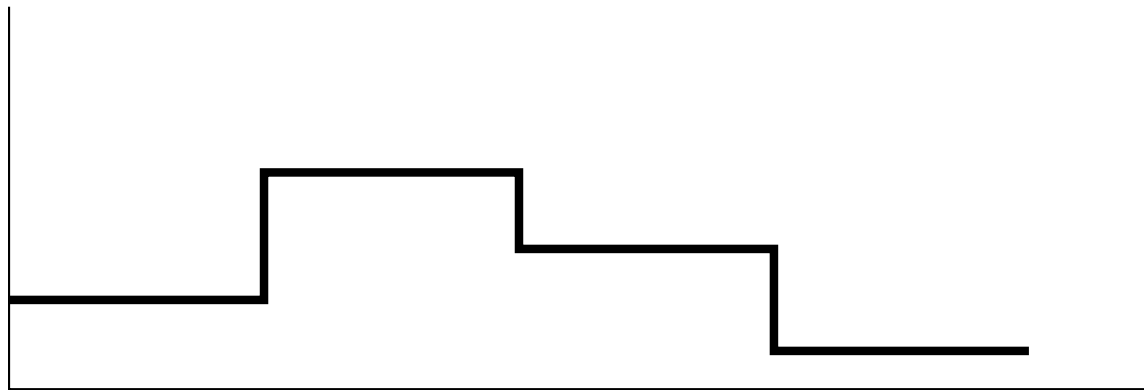
[2 4 8 12 14 0 2 1]



On peut représenter la fonction précédente de manière différente en exploitant une éventuelle corrélation de valeurs voisines.

Pour ce faire, moyennons d'abord les paires de valeurs voisines pour obtenir:

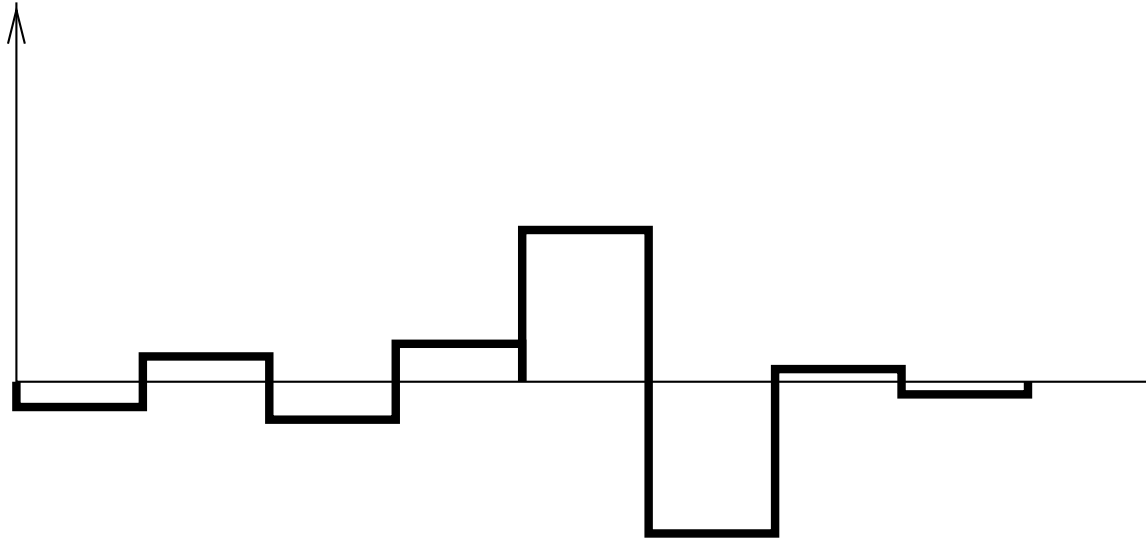
[3 10 7 1.5]



Afin de récupérer le signal initial nous devons également enregistrer d'autres valeurs représentant la perte d'information.

Nous choisirons

$$[-1 \quad -2 \quad 7 \quad 0.5]$$



En effet $3+(-1)=2$, $3-(-1)=4$, $10+(-2)=8$, ...

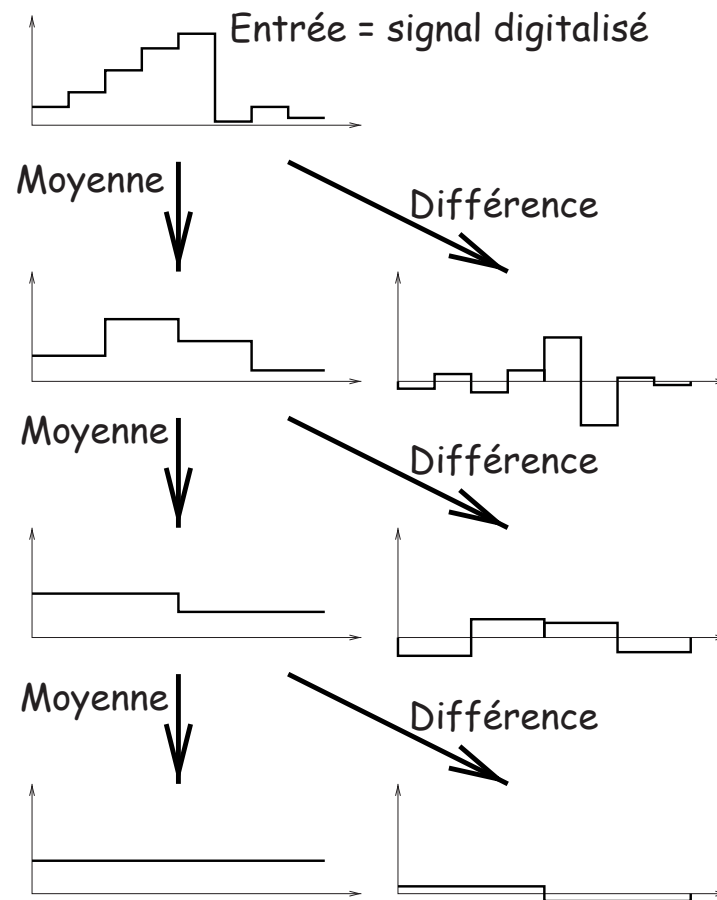
Signal original = signal de résolution plus faible (4 valeurs) et un quadruple de valeurs (les détails).

On peut répéter la procédure précédente sur les moyennes obtenues pour avoir finalement :

Résolution	Moyennes	Détails
8	[2 4 8 12 14 0 2 1]	
4	[3 10 7 1.5]	[-1 -2 7 0.5]
2	[6.5 4.25]	[-3.5 2.75]
1	[5.375]	[1.125]

représentant le signal original par :

[5.375 1.125 -3.5 2.75 -1 -2 7 0.5]



Grand nombre de détails faibles (compression).

On peut considérer la fonction précédente comme une fonction sur $[0,1]$ constante par morceaux sur les intervalles

$$I_{3,k} = [2^{-3}k, 2^{-3}(k+1)[, \quad k = 0, \dots, 2^3 - 1.$$

En notant $\phi(x) = I_{[0,1[}(x)$ et $\phi_{j,k}(x) = \phi(2^j x - k)$, la fonction s'écrit

$$\begin{aligned} f(x) = & 2\phi_{3,0}(x) + 4\phi_{3,1}(x) + 8\phi_{3,2}(x) + 12\phi_{3,3}(x) + \dots \\ & 14\phi_{3,4}(x) + 0\phi_{3,5}(x) + 2\phi_{3,6}(x) + 1\phi_{3,7}(x). \end{aligned}$$

On peut re-écrire alors

$$\begin{aligned} f(x) = & 3\phi_{2,0}(x) + 10\phi_{2,1}(x) + 7\phi_{2,2}(x) + 1.5\phi_{2,3}(x) + \dots \\ & (-1)\psi_{2,0}(x) + (-2)\psi_{2,1}(x) + 7\psi_{2,2}(x) + 0.5\psi_{2,3}(x) \quad , \end{aligned}$$

où

$$\psi(x) = I_{[0,1/2[}(x) - I_{[1/2,1[}(x).$$

Notons :

V_0 le sous-espace vectoriel de $L^2([0, 1[)$ engendré par les fonctions constantes sur $[0, 1[$.

V_1 l'espace vectoriel des fonctions constantes par morceaux sur les intervalles $[0, 1/2[$ et $[1/2, 1[$

V_j l'espace vectoriel des fonctions constantes par morceaux sur les intervalles $I_{j,k}$, $k = 0, 2^j - 1$.

Ils sont emboîtés. De plus pour chacun de ces sous-espaces espaces, les familles $\{\phi_{j,k}, k = 0, \dots, 2^j - 1\}$ forment une base.

Pour le produit usuel $\langle f, g \rangle = \int_0^1 f(x)\bar{g}(x)dx$, ces familles sont orthogonales et la famille $\{\psi_{j,k}, k = 0, \dots, 2^j - 1\}$ est une base de l'espace vectoriel W_j supplémentaire orthogonal de V_j dans V_{j+1} .

Analyse multirésolution sur la droite

Une analyse multirésolution de $L^2(\mathbb{R})$ est une suite emboîtée de sous-espaces fermés $V_j, j \in \mathbb{Z}$, de $L^2(\mathbb{R})$,

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots,$$

telle que

$$\begin{aligned} \bigcap_j V_j &= \{0\}, & \overline{\bigcup_j V_j} &= L^2(\mathbb{R}), \\ f(x) \in V_j &\Leftrightarrow f(2x) \in V_{j+1}, \end{aligned}$$

Il existe une fonction $\phi \in V_0$ telle que

$$V_0 = \left\{ f \in L^2(\mathbb{R}) : f(x) = \sum_{k \in \mathbb{Z}} \alpha_k \phi(x - k) \right\},$$

$\{\phi_k, k \in \mathbb{Z}\}$ est une base "stable" de V_0 , c'est à dire que $0 < m \leq \|\phi_k\| \leq M < \infty$ et $A\|f\|^2 \leq \sum_k \alpha_k^2 \leq B\|f\|^2$.

La fonction ϕ est appelée fonction échelle de l'AMR.

Posons $\phi_{j,k}(x) = 2^{j/2}\phi(2^jx - k)$. Pour une analyse multirésolution orthogonale, une base orthonormée de V_j , est $\{\phi_{j,k} : k \in \mathbb{Z}\}$ et

$$P_j f = \sum_{k \in \mathbb{Z}} \langle f, \phi_{j,k} \rangle \phi_{j,k},$$

est l'approximation de f à la résolution 2^{-j} .

Si W_j est le supplémentaire orthogonal de V_j dans V_{j+1} , on obtient une autre suite $\{W_j : j \in \mathbb{Z}\}$ de sous-espaces fermés orthogonaux de $L^2(\mathbb{R})$, tels que chaque W_j est une dilatation de W_0 , et tels que la somme directe est $L^2(\mathbb{R})$.

L'espace W_j peut éventuellement être choisi comme supplémentaire (non orthogonal) de V_j dans V_{j+1} , conduisant ainsi aux ondelettes bi-orthogonales. Nous ne poursuivrons pas ce point ici.

On peut montrer qu'il existe une fonction ψ telle que l'espace W_0 soit engendré par ses translations entières. On dit que ψ est l'ondelette associée à φ . Dans ce cas il est facile de voir que pour tout entier j , la famille

$$\{\psi_{j,k}, k \in \mathbb{Z}\}$$

est une base orthonormée de W_j .

Si $g \in L^2(\mathbb{R})$ on a :

$$g = \sum_{k \in \mathbb{Z}} c_{j_0,k} \varphi_{j_0,k} + \sum_{j \geq j_0} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k} \quad (1)$$

où j_0 représente un niveau d'approximation grossière. La première partie de (1) est la projection $P_{j_0}g$ de g sur V_{j_0} , et la seconde partie représente les détails.

$$c_{j,k} = \langle g, \varphi_{j,k} \rangle \quad \text{et} \quad d_{j,k} = \langle g, \psi_{j,k} \rangle.$$

Les ondelettes périodiques

Jusqu'à présent nos fonctions étaient définies sur \mathbb{R} .

Pour certaines applications, cela semble raisonnable.

Toutefois, pour des applications en traitement du signal ou de l'image, le domaine sur lequel sont définies les fonctions étudiées est en général borné.

Il existe plusieurs manières de réaliser une analyse multirésolution adaptée à un domaine borné. Nous n'allons introduire ici que le cas des ondelettes périodiques, car c'est la manière la plus intuitive et la plus simple.

Considérons la fonction échelle $\phi \in L^2(\mathbb{R})$ et l'ondelette $\psi \in L^2(\mathbb{R})$ associées à une analyse multirésolution orthogonale de $L^2(\mathbb{R})$.

Pour tout $j, l \in \mathbb{Z}$, on définit la fonction échelle périodique de période 1, par

$$\tilde{\phi}_{j,l}(x) = \sum_{n=-\infty}^{+\infty} \phi_{j,l}(x+n)$$

et l'ondelette périodique associée

$$\tilde{\psi}_{j,l}(x) = \sum_{n=-\infty}^{+\infty} \psi_{j,l}(x+n).$$

On remarquera que, pour $j \leq 0$ et $l \in \mathbb{Z}$,

$$\tilde{\phi}_{j,l}(x) = 2^{j/2} \sum_n \phi(2^j(x+n-2^{-j}l)) = \tilde{\phi}_{j,0}(x)$$

Donc pour $j \leq 0$, $\tilde{\phi}_{j,l}(x) = 2^{-j/2}$. De même on peut montrer que pour tout $j \leq -1$, $\tilde{\psi}_{j,l} = 0$.

Pour une ondelette ϕ et ψ générées par un filtre de longueur D , on a :

$$\tilde{\phi}_{j,l}(x) = 2^{-j/2} \text{ pour tout } j \leq 0 \text{ et tout } l \in \mathbb{Z}.$$

$$\tilde{\psi}_{j,l}(x) = 0 \text{ pour tout } j \leq -1 \text{ et tout } l \in \mathbb{Z}.$$

Pour tout $j > J_0 \geq \log_2(D - 1)$ et $x \in [0, 1]$,

$$\tilde{\phi}_{j,l}(x) = \phi_{j,l}(x)I_{I_{j,l}}(x) + \phi_{j,l}(x + 1)I_{I_{j,l}^c}(x)$$

avec une relation identique pour $\tilde{\psi}$. Pour conséquence, on a :

$$\int_0^1 f(x)\tilde{\phi}_{j,l}(x)dx = \int_{-\infty}^{\infty} \tilde{f}(x)\phi_{j,l}(x)dx$$

(de même pour ψ) où $\tilde{f}(x) = f(x - [x])$, $x \in \mathbb{R}$.

Les fonctions échelles et les ondelettes périodiques génèrent ainsi une analyse multirésolution orthonormée de $L^2([0, 1])$.

Les espaces d'approximation et de détails sont donnés respectivement par

$$\tilde{V}_j = \text{span} \{ \tilde{\phi}_{j,l}, l = 0, \dots, 2^j - 1 \}$$

et

$$\tilde{W}_j = \text{span} \{ \tilde{\psi}_{j,l}, l = 0, \dots, 2^j - 1 \}$$

On obtient ainsi, pour $J_0 \geq 0$ la décomposition de $L^2([0, 1])$:

$$L^2([0, 1]) = \tilde{V}_{J_0} \oplus \left(\bigoplus_{j \geq J_0} \tilde{W}_j \right).$$

Algorithmes rapides de transformation en ondelettes discrètes

L'orthogonalité des fonctions échelles et des ondelettes associées ainsi que le couplage dyadique induit par l'AMR conduisent à un algorithme de calcul rapide pour décomposer ou reconstituer une fonction de V_j .

Considérons $f \in L^2(\mathbb{R})$. On sait que

$$(P_{V_j}f)(x) = (P_{V_{j-1}}f)(x) + (P_{W_{j-1}}f)(x),$$

formulée par

$$(P_{V_j}f)(x) = \sum_{\ell \in \mathbb{Z}} c_{j-1,\ell} \phi_{j-1,\ell}(x) + \sum_{\ell \in \mathbb{Z}} d_{j-1,\ell} \psi_{j-1,\ell}(x).$$

L'objectif est de déduire une relation entre la suite des coefficients $c_{j,\ell}$ et les suites $c_{j-1,\ell}$ et $d_{j-1,\ell}$.

La clé pour cela est de se rappeler que

$$\phi_{j-1,\ell}(x) = \sum_{k=0}^{D-1} a_k \phi_{j,2\ell+k}(x) \text{ et } \psi_{j-1,\ell}(x) = \sum_{k=0}^{D-1} b_k \phi_{j,2\ell+k}(x)$$

On obtient ainsi

$$\begin{aligned} c_{j-1,\ell} &= \int f(x) \phi_{j-1,\ell}(x) dx \\ &= \int f(x) \sum_{k=0}^{D-1} a_k \phi_{j,2\ell+k}(x) dx \\ &= \sum_{k=0}^{D-1} a_k \int f(x) \phi_{j,2\ell+k}(x) dx = \sum_{k=0}^{D-1} a_k c_{j,2\ell+k} \end{aligned}$$

et

$$d_{j-1,\ell} = \sum_{k=0}^{D-1} b_k c_{j,2\ell+k}.$$

Réciproquement, en remarquant encore que

$$\phi_{j,\ell}(x) = \sum_k a_{\ell-2k} \phi_{j-1,k}(x) + \sum_k b_{\ell-2k} \psi_{j-1,k}(x)$$

on obtient

$$c_{j,\ell} = \sum_k a_{\ell-2k} c_{j-1,k} + \sum_k b_{\ell-2k} d_{j-1,k}(x)$$

Les deux équations ci-dessus déterminent la transformée rapide en ondelettes discrètes.

A titre d'illustration, considérons le vecteur

$$\mathbf{c} = \begin{pmatrix} c_{3,0} \\ c_{3,1} \\ c_{3,2} \\ c_{3,3} \\ c_{3,4} \\ c_{3,5} \\ c_{3,6} \\ c_{3,7} \end{pmatrix} .$$

Les étapes de décomposition donnent alors

$$\begin{pmatrix} c_{3,0} \\ c_{3,1} \\ c_{3,2} \\ c_{3,3} \\ c_{3,4} \\ c_{3,5} \\ c_{3,6} \\ c_{3,7} \end{pmatrix} \longrightarrow \begin{pmatrix} c_{2,0} \\ c_{2,1} \\ c_{2,2} \\ \frac{c_{2,3}}{d_{2,0}} \\ d_{2,1} \\ d_{2,3} \\ d_{2,4} \end{pmatrix} \longrightarrow \begin{pmatrix} c_{1,0} \\ \frac{c_{1,1}}{d_{1,0}} \\ d_{1,1} \\ d_{2,0} \\ d_{2,1} \\ d_{2,3} \\ d_{2,4} \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{c_{0,0}}{d_{0,0}} \\ d_{1,0} \\ d_{1,1} \\ d_{2,0} \\ d_{2,1} \\ d_{2,2} \\ d_{2,4} \end{pmatrix} .$$

Le vecteur final est le résultat de la transformation (DWT) des valeurs initiales.

Propriétés d'approximation

Une approximation linéaire d'un signal de carré intégrable f sur une base orthonormée $\mathcal{B} = \{e_n\}_{n \in \mathbb{Z}}$ projette le signal f sur l'espace engendré par M vecteurs choisis a priori dans la \mathcal{B} :

$$f_M = \sum_{n=0}^{M-1} \langle f, e_n \rangle e_n.$$

La qualité de l'approximation

$$\|f - f_M\|_2 = \sum_{n \geq M} |\langle f, e_n \rangle|_2$$

dépend des propriétés de f selon \mathcal{B} .

L'analyse de Fourier fournit des approximations linéaires efficaces pour des signaux uniformément lisses en projetant le signal sur l'espace engendré par des ondes sinusoïdales de fréquences les M premières basses fréquences.

La qualité de l'approximation dépend de la régularité globale de la fonction f .

Dans une base d'ondelettes le signal est projeté sur l'espace V_M .

Là encore, la qualité de l'approximation dépend de la régularité globale de la fonction f .

Approximation non linéaire

L'approximation linéaire de f est améliorée si les M vecteurs sont choisis a posteriori, en dépendant de f .

Pour un nombre M fixé, l'erreur d'approximation est minimisée en prenant les M vecteurs pour lesquels les coefficients $|\langle f, e_n \rangle|$ sont les plus grands.

Si B est une base d'ondelettes, l'amplitude des coefficients est liée à la régularité locale de f et une approximation non linéaire revient à utiliser un pavage adaptatif dont la résolution croît localement là où la fonction est irrégulière.

Première application : Compression

Le but de la compression est d'exprimer un ensemble initial de données en utilisant un ensemble plus petit, avec ou sans perte d'information.

Supposons par exemple que l'on dispose d'une fonction f développée dans une base

$$f(x) = \sum_{k=1}^m c_k e_k(x).$$

Les données initiales dans ce cas sont les coefficients c_1, \dots, c_m . On aimerait alors déterminer une approximation de $f(x)$ exigeant moins de coefficients dans une base éventuellement différente.

Plus précisément, étant donné une borne spécifique d'erreur maximale $\epsilon > 0$ on cherche à déterminer

$$\tilde{f}(x) = \sum_{k=1}^{\tilde{m}} \tilde{c}_k \tilde{e}_k(x)$$

avec $\tilde{m} < m$ et $\|f - \tilde{f}\| \leq \epsilon$.

On peut rechercher une base appropriée $\{\tilde{e}_k\}_{k=1,\dots,\tilde{m}}$.

Pour le moment regardons simplement comment trouver une bonne approximation en utilisant une base fixée à l'avance.

Le problème est assez simple si la base est orthonormée. Soit σ une permutation de $\{1, 2, \dots, m\}$ et soit \tilde{f} la fonction utilisant les \tilde{m} premiers coefficients de la permutation σ :

$$\tilde{f}(x) = \sum_{k=1}^{\tilde{m}} c_{\sigma(k)} e_{\sigma(k)}(x).$$

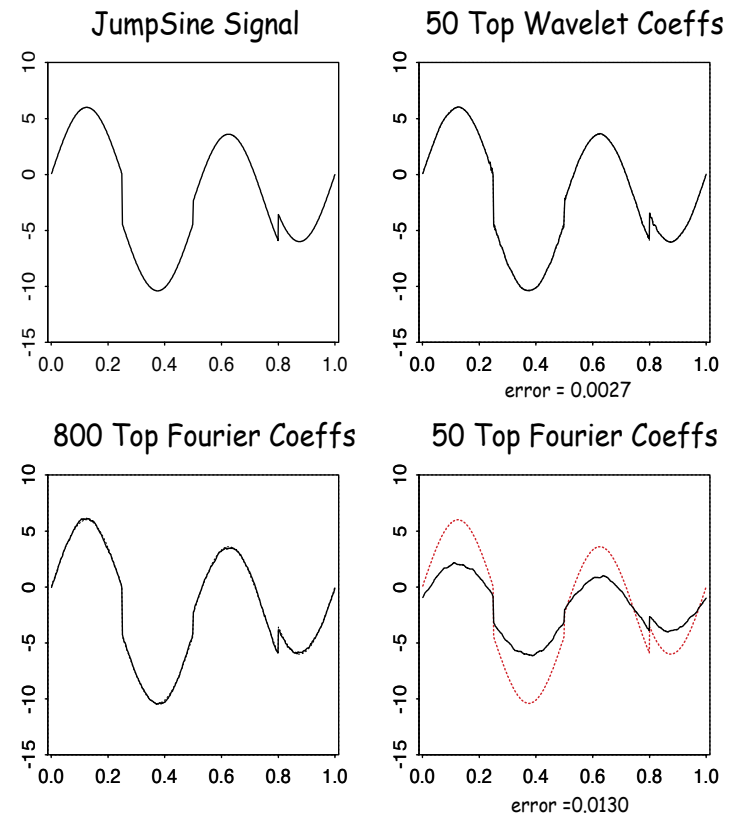
L'erreur L^2 de cette approximation n'est autre que

$$\|f - \tilde{f}\|_2^2 = \sum_{k=\tilde{m}+1}^m |c_{\sigma(k)}|^2$$

Il faut donc pour minimiser cette erreur que σ range les coefficients par ordre décroissant de leur valeur absolue.

Exemple

Compression



Débruitages de signaux

Données :

$$Y_j = f(t_j) + \epsilon_j, \quad j = 1, \dots, n$$

Typiquement : t_j ordonnés et équidistants , $n = 2^m$, et ϵ_j i.i.d de moyenne 0 et de variance σ^2 .

En pratique,

- l'échantillonnage n'est pas équidistant;
- la taille des données n'est pas une puissance de 2;
- l'échantillonnage peut-être aléatoire;
- les erreurs peuvent ne pas être de variance constante;

Méthode linéaire (régularisation)

Objectif : Proposer une méthode d'identification d'une fonction observée avec bruit par une méthode similaire à celle de régularisation par splines de lissage (Laurent (1972), Wahba (1990)) mais plus adaptée à des fonctions moins lisses.

- la méthode est une extension des méthodes de régularisation par pénalisation.
- le choix du paramètre de lissage (ou de pénalisation) par des variantes itératives du principe de validation croisée.
- optimalité théorique et illustrations sur des exemples.

Les splines de lissage

Pour identifier la fonction f à partir des observations bruitées Y_i , les méthodes de régularisation par splines de lissage sont fondées sur la minimisation de la fonctionnelle de coût :

$$\frac{1}{n} \sum_{i=1}^n (Y_i - f(t_i))^2 + \lambda \int_0^1 (f^{(m)}(x))^2 dx,$$

λ étant le paramètre de lissage. La fonction f est supposée appartenir à l'espace de Sobolev W_p^m .

Les ondelettes

L'avantage des ondelettes est qu'elles permettent d'appréhender des fonctions moins lisses et spatialement plus hétérogènes que celles vivant dans des espaces de Sobolev. Pour $f \in L_2([0, 1])$, et $j_0 \in \mathbb{N}$, on a

$$f = \sum_{k=0}^{2^{j_0}-1} \alpha_{j_0,k} \phi_{j_0,k} + \sum_{j \geq j_0} \sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}.$$

Idée simple:

Approximer $2^{m/2} \langle f, \phi_{m,k} \rangle \sim f(k/2^m)$ et remplacer les données brutes $\{Y_i\}$ par son interpolation dans V_m :

$$\hat{f}_m(t) = 2^{-m/2} \sum_{k \in \mathbb{Z}} Y_k \phi_{m,k}(t)$$

Minimiser

$$\|\hat{f}_m - f\|_{L^2([0,1])}^2 + \lambda J_{spp}^p(P_{V_{J_0}} f)$$

où J_0 est donné et J_{spp} est la norme de $B_{pp}^s([0,1])$.

La norme de $g \in B_{pq}^s([0,1])$ est équivalente à

$$J_{spq}(\alpha, \beta) = \|\alpha_{j_0}\|_p + \left(\sum_{j=0}^{\infty} (2^{j(s+(1/2)-(1/p))} \|\beta_j\|_p)^q \right)^{1/q}.$$

La solution

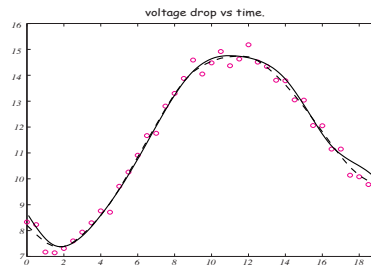
$$f_\lambda = \sum_{k=0}^{2^{j_0}-1} c_{j_0,k} \varphi_{J_0,k} + \sum_{j=J_0}^m \sum_{k=0}^{2^j-1} \hat{\beta}_{j,k} \psi_{j,k},$$

- $c_{j_0,k}$, $k = 0, \dots, 2^{j_0} - 1$ sont les coefficients d'échelle de la DWT de \hat{f}_m .
- $\hat{\beta}_{j,k} = \frac{d_{j,k}}{1+\lambda 2^{2sj}}$, $j \geq j_0, k = 0, \dots, 2^j - 1$, avec $d_{j,k}$ les coefficients d'ondelettes de $W\hat{f}_m$.

Asymptotiquement, quand $N \rightarrow \infty$, si $s > 1/2$ et si $\lambda = \mathcal{O}(N^{-2s/(2s+1)})$, l'erreur quadratique moyenne de f_λ se comporte comme $\mathcal{O}(N^{-2s/(2s+1)})$.

Validation croisée et exemple

On peut choisir J_0 et λ à l'aide des données et montrer que l'estimation qui en résulte possède les bonnes propriétés (au moins asymptotiques). Une des méthodes existantes est la validation croisée.



Ajustement sur des données réelles basé sur des ondelettes de Daubechies d'ordre 8 et comparée au lissage spline.

Méthodes non linéaires

Les méthodes usuelles de débruitage par décompositions en ondelettes reposent sur le fait que pour plusieurs type de signaux leur représentation dans le domaine des coefficients d'ondelettes est creuse, et donc on n'a besoin d'estimer que quelques grands coefficients pour bien estimer la fonction. Le schéma usuel est le suivant :

Développement du signal observé en série d'ondelettes. ■

Extraction des coefficients "significatifs" par méthodes de seuillage ou de rétrécissement. ■

Reconstruction de la fonction par synthèse des coefficients débruités.

Le seuillage ou le rétrécissement permet un bon compromis entre erreur d'approximation et variabilité du résultat, l'erreur globale étant mesurée par une moyenne de la norme L_p , $p \geq 1$ de l'erreur d'identification.

Malgré leur efficacité asymptotique les estimations non linéaires fondées sur les décompositions en ondelettes, les résultats par ce type de méthodes sont souvent moins bons dans le cas d'un nombre modéré d'observations (le seuillage est trop "brutal").

Ceci est principalement dû au fait que l'on ne tient pas compte des informations éventuelles disponibles sur le type de fonctions à débruiter.

L'ensemble des méthodes non-linéaires existantes sont du type :

$$2^{-1} \sum_{i=1}^n (z_i - \theta_i)^2 + \lambda \sum_{i \geq i_0} p(|\theta_i|), \quad (2)$$

où z_i est la i^e ligne de $\mathbf{z} = \mathbf{W}\mathbf{Y}_n$ et p_λ est une fonction de pénalité appropriée.

Seuillage dur

L'idée du seuillage dur :

1. Transformer les données bruitées via DWT : $\tilde{\Theta} = W \cdot Y$.
2. Pour séparer le signal du bruit, seuiller les coefficients: poser $\hat{\theta}_{j,k} = \tilde{\theta}_{j,k}$ si $|\tilde{\theta}_{j,k}|$ est grand, 0 sinon.
3. Estimer le signal par $\hat{f} = W^{-1} \cdot \hat{\Theta}$.

Seuillage doux

1. Transformer les données bruitées via DWT : $\tilde{\Theta} = W \cdot Y$.
2. Pour séparer le signal du bruit, seuiller les coefficients: poser

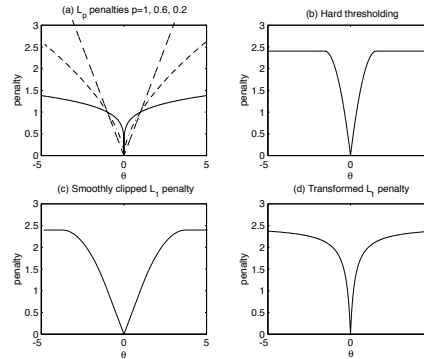
$$\hat{\theta}_{j,k} = \begin{cases} \tilde{\theta}_{j,k} - \lambda, & \text{si } |\tilde{\theta}_{j,k}| > \lambda, \\ 0, & \text{si } |\tilde{\theta}_{j,k}| \leq \lambda, \end{cases}$$

3. Estimer le signal par $\hat{f} = W^{-1} \cdot \hat{\Theta}$.

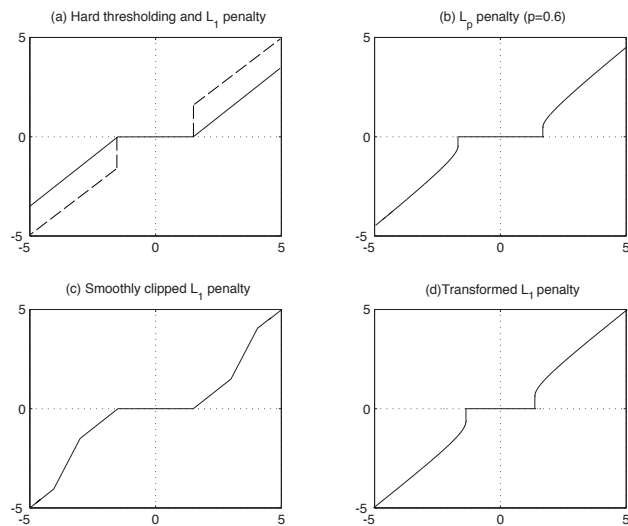
Il existe plusieurs manières pour choisir les seuils. La résolution initiale j_0 est habituellement prise égale à $(\log_2 N)/2$.

Exemples de fonctions de seuillage

(a) Pénalité L_p avec $p = 1$ (seuillage doux ($p=1$)), $p = 0.6$ (short dash) et $p = 0.2$ (solide); (b) seuillage dur; (c) SCAD (robuste); (d) Seuillage doux transformé.

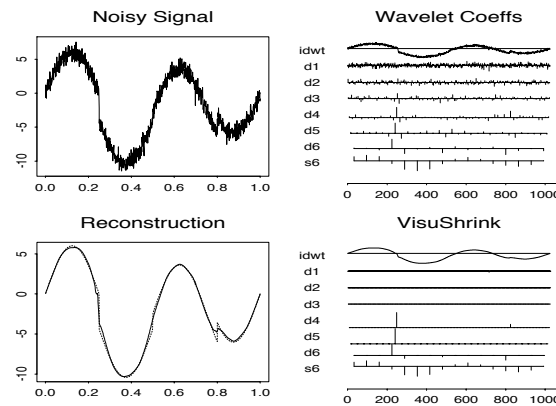


Les estimateurs correspondants :



Seuillage Universel (VisuShrink)

Lorsque le bruit est gaussien, la plupart des coefficients (normalisés selon: $\sqrt{N}d_{j,k}/\sigma$) sont essentiellement du bruit blanc. Cela suggère comme seuil $\lambda = \sqrt{2 \log N}$, produisant le principe VisuShrink implémenté dans Wavelab.



Seuillage minimax

Les performances de l'estimateur sont comparées à

$$B_n(\boldsymbol{\theta}, \sigma^2) = \sigma^2 + \sum_{i=1}^n \min(\theta_i, \sigma^2)$$

où $\boldsymbol{\theta}$ est le vecteur n -dimensionnel des coefficients d'ondelette, bruités par un bruit blanc de variance σ^2 .

Un des résultats les plus significatifs de Donoho et Johnstone est que , si l'on observe un vecteur gaussien n -dimensionnel $\mathbf{U} \sim \mathbf{N}_n(\boldsymbol{\theta}, \sigma^2 \mathbf{I}_n)$,

$$\sup_{\boldsymbol{\theta} \in \mathbb{R}^n} \frac{\mathbb{E} \|\eta_{\lambda_n}^S(\mathbf{U}) - \boldsymbol{\theta}\|^2}{B_n(\boldsymbol{\theta}, \sigma^2)} \leq (1 + 2 \log n),$$

quand $\lambda_n = \sqrt{2 \log n}$. Il est montré de plus que le facteur $2 \log n$ ne peut être diminué.

Les seuils optimaux λ_n^* n'existent pas sous forme analytique mais des approximations numériques pour diverses valeurs de n sont disponibles dans Wavelab.

Performances pour les petits échantillons

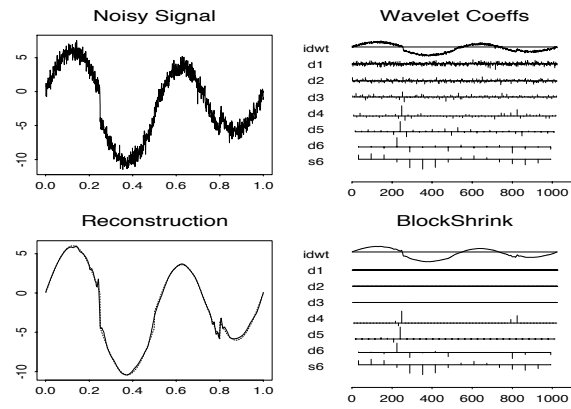
Les résultats par ce type de méthodes sont souvent moins bons dans le cas d'un nombre modéré d'observations et le type de seuillage à adopter dans ce cas peut être important.

Lorsque l'on compare le seuillage doux et le seuillage dur, le dur tend à avoir une variance plus importante que le doux, alors que ce dernier présente un biais plus important.

Certains auteurs proposent quelque chose de plus robuste :

$$\hat{\theta}_{j,k}^{\lambda_1, \lambda_2} = \begin{cases} 0 & \text{si } |\tilde{\theta}_{j,k}| \leq \lambda_1 \\ \text{sgn}(\tilde{\theta}_{j,k}) \frac{\lambda_2 (|\tilde{\theta}_{j,k}| - \lambda_1)}{\lambda_2 - \lambda_1} & \text{si } \lambda_1 < |\tilde{\theta}_{j,k}| \leq \lambda_2 \\ \tilde{\theta}_{j,k} & \text{si } |\tilde{\theta}_{j,k}| > \lambda_2, \end{cases}$$

offrant les avantages des deux types de seuillages.



Problèmes et extensions

A Pénalité non séparables

B Choix de λ (minimiser une borne supérieure du risque? utiliser la validation-croisée? des procédures de sélection de modèles?)

C Autres problèmes d'estimation fonctionnelle.

D Exploiter les liens entre modèles bayésiens d'ondelettes utilisés en analyse d'image.

Tous les algorithmes sont implémentés en MATLAB.

Disponibles à

www-lmc.imag.fr/SMS/software