Ondelettes - Applications statistiques

Stéphan Clémençon

Telecom ParisTech

April 2, 2014

Agenda

- **Problèmes statistiques**

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs

Agenda

- **Problèmes statistiques**
    - régression - débruitage - filtrage
    - Estimation d'une densité de probabilité
    - Problèmes inverses mal posés - mesures indirectes
    - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
    - Analyse harmonique computationnelle
    - Types de seuillage: dur, doux, par blocs

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence
- **Choix d'une "meilleure" base/représentation**

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence
- **Choix d'une "meilleure" base/représentation**
  - Poursuite de base

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence
- **Choix d'une "meilleure" base/représentation**
  - Poursuite de base
  - Minimisation de l'entropie - l'algorithme de Coifman-Wickerhauser

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence
- **Choix d'une "meilleure" base/représentation**
  - Poursuite de base
  - Minimisation de l'entropie - l'algorithme de Coifman-Wickerhauser
  - "Matching Pursuit"

Agenda

- **Problèmes statistiques**
  - régression - débruitage - filtrage
  - Estimation d'une densité de probabilité
  - Problèmes inverses mal posés - mesures indirectes
  - Estimation d'une fonction de covariance
- **Retour sur le "wavelet shrinkage"**
  - Analyse harmonique computationnelle
  - Types de seuillage: dur, doux, par blocs
- **Ondelettes-vaguelettes: mesures indirectes**
- **Paquets d'ondelettes et paquets de cosinus locaux**
  - Pavage du plan temps-fréquence
  - Structure arborescente
  - Dictionnaire d'atomes temps-fréquence
- **Choix d'une "meilleure" base/représentation**
  - Poursuite de base
  - Minimisation de l'entropie - l'algorithme de Coifman-Wickerhauser
  - "Matching Pursuit"
- **Au delà des ondelettes:** ridgelets, edgelets, curvelets, *etc.*

Problèmes statistiques - Débruitage

- On observe

$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

Problèmes statistiques - Débruitage

- On observe
$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \ \ 1 \leq i \leq n$$

- But: recouvrer le signal $f(t)$ à partir des observations y_i

Problèmes statistiques - Débruitage

- On observe

$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \ \ 1 \leq i \leq n$$

- But: recouvrer le signal $f(t)$ à partir des observations y_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(t) - f(t)\right)^2\right]$$

Problèmes statistiques - Débruitage

- On observe

$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer le signal $f(t)$ à partir des observations y_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(t) - f(t)\right)^2\right]$$

- Hypothèse: les ϵ_i sont centrés, i.i.d. de loi $\mathcal{N}(0, 1)$

Heuristique:

Lisser pour supprimer les variations haute-fréquence *i.e.* le bruit

Problèmes statistiques - Débruitage

- On observe

$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer le signal $f(t)$ à partir des observations y_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(t) - f(t)\right)^2\right]$$

- Hypothèse: les ϵ_i sont centrés, i.i.d. de loi $\mathcal{N}(0, 1)$

Heuristique:

Lisser pour supprimer les variations haute-fréquence *i.e.* le bruit

- Choix de la fenêtre de lissage? du noyau (régularité)?

Problèmes statistiques - Débruitage

- On observe

$$y_i = f\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer le signal $f(t)$ à partir des observations y_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(t) - f(t)\right)^2\right]$$

- Hypothèse: les ϵ_i sont centrés, i.i.d. de loi $\mathcal{N}(0, 1)$

Heuristique:

Lisser pour supprimer les variations haute-fréquence *i.e.* le bruit

- Choix de la fenêtre de lissage? du noyau (régularité)?
- En pratique, la régularité de f est inconnue, σ aussi

Problèmes statistiques - Estimation de densité

- On observe des données rélles:

$$X_1,\ \ldots,\ X_n \overset{i.i.d.}{\sim} F(x) = f(x)dx$$

Problèmes statistiques - Estimation de densité

- On observe des données rélles:

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} F(x) = f(x)dx$$

- But: recouvrer la densité $f(x)$ à partir des observations X_i

Construire $\widehat{f}(x)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right]$$

avec $1 \leq p \leq \infty$.

Problèmes statistiques - Estimation de densité

- On observe des données rélles:

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} F(x) = f(x)dx$$

- But: recouvrer la densité $f(x)$ à partir des observations X_i

Construire $\widehat{f}(x)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right]$$

avec $1 \leq p \leq \infty$.

Heuristique:

Régulariser la distribution empirique $\widehat{F}_n(dx) = n^{-1}\sum_{i=1}^n \delta_{X_i}$

Problèmes statistiques - Estimation de densité

- On observe des données rélles:

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} F(x) = f(x)dx$$

- But: recouvrer la densité $f(x)$ à partir des observations X_i

Construire $\widehat{f}(x)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right]$$

avec $1 \leq p \leq \infty$.

Heuristique:

Régulariser la distribution empirique $\widehat{F}_n(dx) = n^{-1}\sum_{i=1}^n \delta_{X_i}$

- Choix de la fenêtre de lissage? du noyau (régularité)?

Problèmes statistiques - Estimation de densité

- On observe des données rélles:

$$X_1, \ldots, X_n \overset{i.i.d.}{\sim} F(x) = f(x)dx$$

- But: recouvrir la densité $f(x)$ à partir des observations X_i

Construire $\widehat{f}(x)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right]$$

avec $1 \leq p \leq \infty$.

Heuristique:

Régulariser la distribution empirique $\widehat{F}_n(dx) = n^{-1} \sum_{i=1}^n \delta_{X_i}$

- Choix de la fenêtre de lissage? du noyau (régularité)?
- En pratique, la régularité de f est inconnue

- On dispose de *mesures indirectes* bruitées:

$$y_i = Kf\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \;\; 1 \leq i \leq n$$

- On dispose de *mesures indirectes* bruitées:

$$y_i = Kf\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer la fonction $f(t)$ à partir des observations X_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right].$$

Problèmes statistiques - Problèmes inverses mal-posés

- On dispose de *mesures indirectes* bruitées:

$$y_i = Kf\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer la fonction $f(t)$ à partir des observations X_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right].$$

Heuristique:

Pour de nombreux opérateurs, les ondelettes sont "presque" des fonctions propres.

Problèmes statistiques - Problèmes inverses mal-posés

- On dispose de *mesures indirectes* bruitées:

$$y_i = Kf\left(\frac{i}{n}\right) + \sigma \cdot \epsilon_i, \quad 1 \leq i \leq n$$

- But: recouvrer la fonction $f(t)$ à partir des observations X_i

Construire $\widehat{f}(t)$ de façon à minimiser:

$$R_p(f, \widehat{f}) = \mathbb{E}\left[\left(\widehat{f}(x) - f(x)\right)^p\right].$$

Heuristique:

Pour de nombreux opérateurs, les ondelettes sont "presque" des fonctions propres.

- "Quasi-diagonaliser" l'opérateur K - Alternative à la SVD

Problèmes statistiques - Estimation de la Covariance

- On observe un processus $X_1, \ldots, X_n$ du second ordre

Problèmes statistiques - Estimation de la Covariance

- On observe un processus $X_1, \ldots, X_n$ du second ordre
- But: estimer la fonction de covariance

$$\Gamma(k, l) = cov(X_k, X_l)$$

Problèmes statistiques - Estimation de la Covariance

- On observe un processus $X_1, \ldots, X_n$ du second ordre
- But: estimer la fonction de covariance

$$\Gamma(k,l) = cov(X_k, X_l)$$

- Cas stationnaire: $\Gamma(k,l) = \gamma(k-l)$

$$||\widehat{\gamma} - \gamma||_{l^2(\mathbb{N})} = O(n^{-1})$$

Problèmes statistiques - Estimation de la Covariance

- On observe un processus $X_1, \ldots, X_n$ du second ordre
- But: estimer la fonction de covariance

$$\Gamma(k,l) = cov(X_k, X_l)$$

- Cas stationnaire: $\Gamma(k,l) = \gamma(k-l)$

$$||\widehat{\gamma} - \gamma||_{l^2(\mathbb{N})} = O(n^{-1})$$

- Cas non-stationnaire: en général,

$$||\widehat{\Gamma} - \Gamma||_{HS} = O(n)$$

où $\widehat{\Gamma}(k,l) = X_K X_l$ désigne la covariance empirique et $||M||_{HS}^2 = \sum_{i,j} m_{i,j}^2 = tr(MM^*)$.

Problèmes statistiques - Estimation de la Covariance

- Estimation dans la base canonique

Théorème

Si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien, alors

$$\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] = ||\Gamma||_{HS}^2 + \left(\mathbb{E}||X||^2\right)^2.$$

Problèmes statistiques - Estimation de la Covariance

- Estimation dans la base canonique

Théorème

Si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien, alors

$$\mathbb{E}[||\widehat{\Gamma} - \Gamma||^2_{HS}] = ||\Gamma||^2_{HS} + \left(\mathbb{E}||X||^2\right)^2.$$

- Si $var(X_i) = \sigma^2$, alors le second terme est de l'ordre $\sim \sigma^4 n^2$

Problèmes statistiques - Estimation de la Covariance

- Estimation dans la base canonique

Théorème

Si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien, alors

$$\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] = ||\Gamma||_{HS}^2 + \left(\mathbb{E}||X||^2\right)^2.$$

- Si $var(X_i) = \sigma^2$, alors le second terme est de l'ordre $\sim \sigma^4 n^2$
- Estimation dans une b.o.n. diagonalisante B (base de Karhunen-Loeve): il suffit d'estimer les n valeurs propres

$$\begin{aligned} \mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] &= \mathbb{E}[||B^* \widehat{\Gamma} B - D||_{HS}^2] \\ &\sim n \end{aligned}$$

Problèmes statistiques - Estimation de la Covariance

- Estimation dans la base canonique

Théorème

Si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien, alors

$$\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] = ||\Gamma||_{HS}^2 + \left(\mathbb{E}||X||^2\right)^2.$$

- Si $var(X_i) = \sigma^2$, alors le second terme est de l'ordre $\sim \sigma^4 n^2$
- Estimation dans une b.o.n. diagonalisante B (base de Karhunen-Loeve): il suffit d'estimer les n valeurs propres

$$\begin{aligned}\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] &= \mathbb{E}[||B^*\widehat{\Gamma}B - D||_{HS}^2] \\ &\sim n\end{aligned}$$

- Dans le cas stationnaire, la base de Fourier (circulaire) est diagonalisante

Problèmes statistiques - Estimation de la Covariance

- Estimation dans la base canonique

Théorème

Si $X = (X_1, \ldots, X_n)$ est un vecteur gaussien, alors

$$\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] = ||\Gamma||_{HS}^2 + \left(\mathbb{E}||X||^2\right)^2.$$

- Si $var(X_i) = \sigma^2$, alors le second terme est de l'ordre $\sim \sigma^4 n^2$
- Estimation dans une b.o.n. diagonalisante B (base de Karhunen-Loeve): il suffit d'estimer les n valeurs propres

$$\begin{aligned}\mathbb{E}[||\widehat{\Gamma} - \Gamma||_{HS}^2] &= \mathbb{E}[||B^*\widehat{\Gamma}B - D||_{HS}^2] \\ &\sim n\end{aligned}$$

- Dans le cas stationnaire, la base de Fourier (circulaire) est diagonalisante
- Dans la cas localement stationnaire?

Régression/débruitage avec design aléatoire

- Modèle statistique:

$$Y_i = f(X_i) + \epsilon_i, \quad 1 \leq i \leq n,$$

avec $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ conditionnellement à la suite (X_n) et $X_i \overset{i.i.d.}{\sim} \mu$

Régression/débruitage avec design aléatoire

- Modèle statistique:

$$Y_i = f(X_i) + \epsilon_i, \ \ 1 \leq i \leq n,$$

  avec $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ conditionnellement à la suite (X_n) et $X_i \overset{i.i.d.}{\sim} \mu$
- Peut-on transposer l'approche par ondelettes à ce cadre?

Régression/débruitage avec design aléatoire

- Modèle statistique:

$$Y_i = f(X_i) + \epsilon_i, \;\; 1 \leq i \leq n,$$

  avec $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ conditionnellement à la suite (X_n) et $X_i \overset{i.i.d.}{\sim} \mu$
- Peut-on transposer l'approche par ondelettes à ce cadre?
- Oui, si $\mu = \mathcal{U}([0,1])$ car

$$\widehat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} Y_i \psi_{j,k}(X_i)$$

  est alors un estimateur consistant et sans biais de de $\beta_{j,k} = \langle f, \psi_{j,k} \rangle$

Régression/débruitage avec design aléatoire

- Modèle statistique:

$$Y_i = f(X_i) + \epsilon_i, \;\; 1 \leq i \leq n,$$

avec $\epsilon_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ conditionnellement à la suite (X_n) et $X_i \overset{i.i.d.}{\sim} \mu$

- Peut-on transposer l'approche par ondelettes à ce cadre?
- Oui, si $\mu = \mathcal{U}([0,1])$ car

$$\widehat{\beta}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} Y_i \psi_{j,k}(X_i)$$

est alors un estimateur consistant et sans biais de de $\beta_{j,k} = \langle f, \psi_{j,k} \rangle$

- Cas général: $\mathbb{E}[\widehat{\beta}_{j,k}] = \int \psi_{j,k}(x)\mu(dx) \neq \beta_{j,k}$
- La famille $\{\psi_{j,k}\}$ n'est pas une b.o.n. de $L_2(\mu)$ en général.

Régression/débruitage avec design aléatoire

- Idée: déformer les ondelettes: $\mu(dx) = g(x)dx$, si $G(x) = \int_{-\infty}^{x} g(u)du$ était connue,

$$\widetilde{\beta}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} \psi_{j,k}(G(X_i))Y_i$$

  serait un bon estimateur de $\beta_{j,k}$

- Remplacer G par sa contrepartie empirique

$$\widehat{G}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{X_i \leq x\},$$

  $\widetilde{\beta}_{j,k} \to \bar{\beta}_{j,k}$ et considérer l'estimateur

$$\widetilde{f}(x) = \sum_{(j,k)} \bar{\beta}_{j,k}\mathbb{I}\{|\bar{\beta}_{j,k}| \geq \tau\}\psi_{j,k}(\widehat{G}_n(x))$$

Retour sur le "Wavelet shrinkage"

- Soit $f(t) \in L_p$, on peut décomposer le signal selon

$$
\begin{aligned}
f(t) &= \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t) + \sum_{j \geq j_1} \beta_{j,k}\psi_{jk}(t) \\
&= \lim_{j_1 \to \infty} \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t).
\end{aligned}
$$

Retour sur le "Wavelet shrinkage"

- Soit $f(t) \in L_p$, on peut décomposer le signal selon

$$\begin{aligned} f(t) &= \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t) + \sum_{j \geq j_1} \beta_{j,k}\psi_{jk}(t) \\ &= \lim_{j_1 \to \infty} \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t). \end{aligned}$$

Reconstruction sélective:

Ne conserver que les détails aux échelles $j_1 \leq j \leq j_0$ **significatifs**, *i.e.* supérieurs à un certain seuil:

$$\tilde{\beta}_{jk} = \beta_{jk}\mathbb{I}\{|\beta_{jk}| \geq T_j\}.$$

Retour sur le "Wavelet shrinkage"

- Soit $f(t) \in L_p$, on peut décomposer le signal selon

$$
\begin{aligned}
f(t) &= \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t) + \sum_{j \geq j_1} \beta_{j,k}\psi_{jk}(t) \\
&= \lim_{j_1 \to \infty} \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t).
\end{aligned}
$$

Reconstruction sélective:

Ne conserver que les détails aux échelles $j_1 \leq j \leq j_0$ **significatifs**, *i.e.* supérieurs à un certain seuil:

$$\tilde{\beta}_{jk} = \beta_{jk}\mathbb{I}\{|\beta_{jk}| \geq T_j\}.$$

- Approximation non linéaire par seuillage d'ondelettes

$$WS(f) = \sum_k \alpha_{j_1,k}\phi_{j_1,k}(t) + \sum_{j_1 \leq j \leq j_0} \tilde{\beta}_{j,k}\psi_{jk}(t)$$

Approximation linéaire *vs.* approximation non linéaire

- Contrôle de l'erreur d'approximation:

$$
\begin{aligned}
f - WS(f) &= \{f - P_{j_0} f\} + \sum_{j_1 \le j \le j_0} \sum_k \beta_{j,k} \mathbb{I}\{|\beta_{jk}| \le \tau_j\} \psi_{jk} \\
&= \text{linéaire} + \text{non linéaire}
\end{aligned}
$$

avec $P_{j_0} f = \sum_k \alpha_{j_0,k} \phi_{j_0,k}$.

Approximation linéaire *vs.* approximation non linéaire

- Contrôle de l'erreur d'approximation:

$$
\begin{aligned}
f - WS(f) &= \{f - P_{j_0} f\} + \sum_{j_1 \leq j \leq j_0} \sum_k \beta_{j,k} \mathbb{I}\{|\beta_{jk}| \leq \tau_j\} \psi_{jk} \\
&= \text{linéaire} + \text{non linéaire}
\end{aligned}
$$

avec $P_{j_0} f = \sum_k \alpha_{j_0,k} \phi_{j_0,k}$.

- Hypothèse "Besov" $||f||_{\sigma pq} \leq R < \infty$, avec

$$
\begin{aligned}
||f||_{\sigma pq} &= \left(\sum_k |\alpha_{j_1,k}|^p \right)^{1/p} \\
&+ \left(\sum_{j=j_1}^{\infty} 2^{j(\sigma+1/2-1/p)q} \left(\sum_k |\beta_{j,k}|^p \right)^{q/p} \right)^{1/q}
\end{aligned}
$$

Approximation linéaire *vs.* approximation non linéaire

- Approximation linéaire:

$$||f - \sum_k \alpha_{j,k}\phi_{j,k}||_{p'} \leq C(R) \cdot 2^{-j_0(\sigma - 1/p + 1/p')}$$

Approximation linéaire *vs.* approximation non linéaire

- Approximation linéaire:

$$||f - \sum_k \alpha_{j,k}\phi_{j,k}||_{p'} \leq C(R) \cdot 2^{-j_0(\sigma - 1/p + 1/p')}$$

- **"Wavelet thresholding":** le terme résiduel non linéaire $\sum_{j_1 \leq j \leq j_0} || \sum_k \beta_{j,k} \mathbb{I}\{|\beta_{j,k}| \leq \tau_j\} \psi_{jk}(t)||_{p'}$ est majoré par

$$\left(\sum_{j_1 \leq j \leq j_0} \left(2^{j\epsilon/p'} \tau_j^{1-p/p'} \right)^2 \right)^{1/2},$$

avec $\epsilon = \sigma_p - (p' - p)/2$

Approximation linéaire *vs.* approximation non linéaire

- On pose

$$v = \min\left\{\frac{\sigma}{\sigma+1},\ \frac{\sigma - 1/p + 1/p'}{1+2(\sigma - 1/p)}\right\}$$

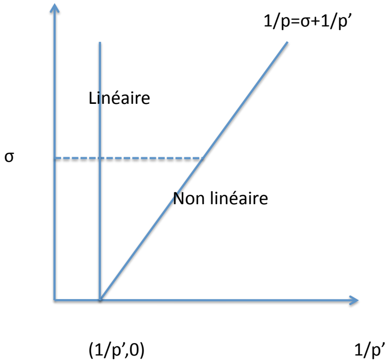
  On a: $v = \sigma/(1+\sigma) \Leftrightarrow \epsilon \geq 0$

Approximation linéaire *vs.* approximation non linéaire

- On pose
$$v = \min\left\{\frac{\sigma}{\sigma+1},\ \frac{\sigma - 1/p + 1/p'}{1+2(\sigma - 1/p)}\right\}$$
On a: $v = \sigma/(1+\sigma) \Leftrightarrow \epsilon \geq 0$
- Lorsque $p' > p$, on choisit:
$$\begin{array}{ccc} \tau_j & \sim & \sqrt{j/n} \\ 2^{j_1} & \sim & n^{1-2v} \\ 2^{j_0} & \sim & n^{v/(\sigma-1/p+1/p')} \end{array}$$

Grâce au seuillage, n^{1-2v} coefficients permettent d'obtenir une erreur d'approximation d'ordre n^{-v}

Approximation linéaire *vs.* approximation non linéaire



Seuillage des coefficients d'ondelettes

- Seuillage "doux" :

$$(\beta_{j,k} - sign(\beta_{j,k})\gamma) \cdot \mathbb{I}\{|\beta_{j,k}| > \gamma\} \rightarrow \beta_{j,k}$$

Seuillage des coefficients d'ondelettes

- Seuillage "doux":

$$(\beta_{j,k} - sign(\beta_{j,k})\gamma) \cdot \mathbb{I}\{|\beta_{j,k}| > \gamma\} \rightarrow \beta_{j,k}$$

- Seuillage fort:

$$\beta_{j,k} \cdot \mathbb{I}\{|\beta_{j,k}| > \gamma\} \rightarrow \beta_{j,k}$$

Seuillage des coefficients d'ondelettes

- Seuillage "doux":

$$(\beta_{j,k} - sign(\beta_{j,k})\gamma) \cdot \mathbb{I}\{|\beta_{j,k}| > \gamma\} \rightarrow \beta_{j,k}$$

- Seuillage fort:

$$\beta_{j,k} \cdot \mathbb{I}\{|\beta_{j,k}| > \gamma\} \rightarrow \beta_{j,k}$$

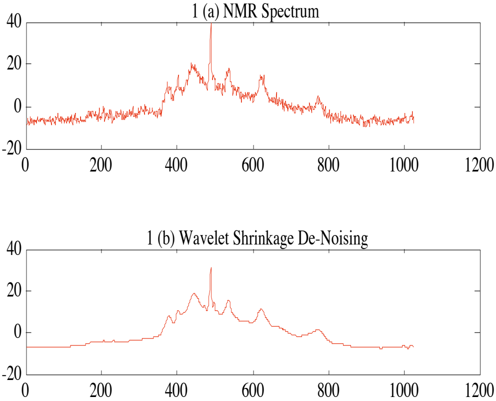
- Seuillage par blocs:

$$\beta_{j,k} \cdot \mathbb{I}\{|b_{j,k}| > \gamma\} \rightarrow \beta_{j,k},$$

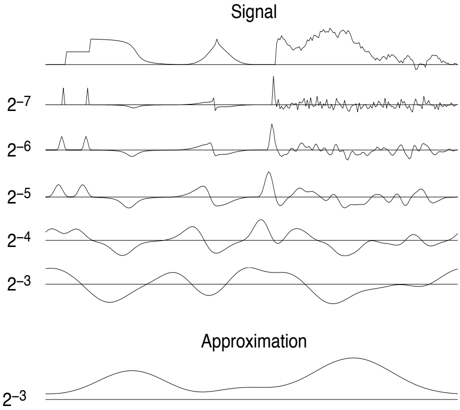
where

$$b_{j,k} = \left(L^{-1} \sum_{k \in B_{j,K}} |\beta_{j,k}|^p \right)^{1/p}$$

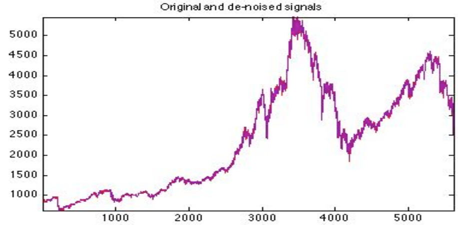
Seuillage des coefficients d'ondelettes

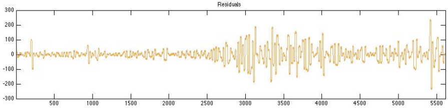


Seuillage des coefficients d'ondelettes



Seuillage des coefficients d'ondelettes





Mesures indirectes

- Modèle:

$$Y_i = Kf(x_i) + \epsilon_i, \quad 1 \leq i \leq n,$$

avec $\{\epsilon_i\}$ bruit blanc gaussien

Mesures indirectes

- Modèle:

$$Y_i = Kf(x_i) + \epsilon_i, \ \ 1 \leq i \leq n,$$

  avec $\{\epsilon_i\}$ bruit blanc gaussien

- Exemples:
  - IRM: K est la transformée de Fourier

Mesures indirectes

- Modèle:
$$Y_i = Kf(x_i) + \epsilon_i,\ \ 1 \leq i \leq n,$$
avec $\{\epsilon_i\}$ bruit blanc gaussien

- Exemples:
  - IRM: K est la transformée de Fourier
  - Spectroscopie de fluorescence : K est la transformée de Laplace

Mesures indirectes

- Modèle:

$$Y_i = Kf(x_i) + \epsilon_i, \;\; 1 \leq i \leq n,$$

  avec $\{\epsilon_i\}$ bruit blanc gaussien

- Exemples:
  - IRM: K est la transformée de Fourier
  - Spectroscopie de fluorescence : K est la transformée de Laplace
  - Imagerie médicale: transformée de Radon

Mesures indirectes

- Modèle:

$$Y_i = Kf(x_i) + \epsilon_i, \quad 1 \leq i \leq n,$$

  avec $\{\epsilon_i\}$ bruit blanc gaussien

- Exemples:
  - IRM: K est la transformée de Fourier
  - Spectroscopie de fluorescence : K est la transformée de Laplace
  - Imagerie médicale: transformée de Radon
  - K est un opérateur de convolution dans de nombreuses applications

Mesures indirectes

- Décomposition en valeurs singulières (SVD):

Mesures indirectes

- Décomposition en valeurs singulières (SVD):
  - diagonaliser l'opérateur autoadjoint K^*K dans une b.o.n. $\{e_n\}$ soit $\{\lambda_n^2\} = Spec(K^*K)$ et $K^*Ke_n = \lambda_n^2 e_n$

$$f = \sum_n \lambda_n^{-1} \langle Kf, h_n \rangle e_n,$$

avec $h_n = ||Ke_n||^{-1} Ke_n$

Mesures indirectes

- Décomposition en valeurs singulières (SVD):
  - diagonaliser l'opérateur autoadjoint K^*K dans une b.o.n. $\{e_n\}$ soit $\{\lambda_n^2\} = Spec(K^*K)$ et $K^*Ke_n = \lambda_n^2 e_n$

$$f = \sum_n \lambda_n^{-1} \langle Kf, h_n \rangle e_n,$$

    avec $h_n = ||Ke_n||^{-1} Ke_n$
  - Régularisation de Tikhonov: $\lambda_n/(\kappa^2 + \lambda_n^2) \to 1/\lambda_n$

Mesures indirectes

- Décomposition en valeurs singulières (SVD):
  - diagonaliser l'opérateur autoadjoint K^*K dans une b.o.n. $\{e_n\}$ soit $\{\lambda_n^2\} = Spec(K^*K)$ et $K^*Ke_n = \lambda_n^2 e_n$

$$f = \sum_n \lambda_n^{-1} \langle Kf, h_n \rangle e_n,$$

    avec $h_n = ||Ke_n||^{-1} Ke_n$
  - Régularisation de Tikhonov: $\lambda_n/(\kappa^2 + \lambda_n^2) \to 1/\lambda_n$
  - Procédure statistique: estimer un nombre fini de coefficients $\langle Kf, h_n \rangle$ et construire un estimateur de f à partir d'une reconstruction tronquée

- Décomposition en valeurs singulières (SVD):
  - diagonaliser l'opérateur autoadjoint K^*K dans une b.o.n. $\{e_n\}$ soit $\{\lambda_n^2\} = Spec(K^*K)$ et $K^*Ke_n = \lambda_n^2 e_n$

$$f = \sum_n \lambda_n^{-1} \langle Kf, h_n \rangle e_n,$$

    avec $h_n = ||Ke_n||^{-1} Ke_n$

  - Régularisation de Tikhonov: $\lambda_n/(\kappa^2 + \lambda_n^2) \to 1/\lambda_n$
  - Procédure statistique: estimer un nombre fini de coefficients $\langle Kf, h_n \rangle$ et construire un estimateur de f à partir d'une reconstruction tronquée
  - Il n'y a a priori aucune raison pour que la base $\{e_n\}$ permette de représenter efficacement la fonction f!

Décomposition en Ondelettes/Vaguelettes

- Les ondelettes sont des bases quasi-diagonalisantes de nombreux opérateurs

Décomposition en Ondelettes/Vaguelettes

- Les ondelettes sont des bases quasi-diagonalisantes de nombreux opérateurs
- Si $\kappa_\gamma \in \mathcal{D}(K^*)$ t.q. $K^*\kappa_\gamma = \psi_\gamma$:

$$\begin{aligned} f &= \sum_\gamma \langle f, \psi_\gamma \rangle \psi_\gamma \\ &= \sum_\gamma \langle Kf, \kappa_\gamma \rangle \psi_\gamma \end{aligned}$$

- Hypothèse: $\exists \{\lambda_j\}$ t.q. si $u_{j,k} = \lambda_j \cdot \kappa_{j,k}$, alors

$$|| \sum_\gamma \alpha_\gamma u_\gamma ||_{L_2} \leq C \left(\sum_\gamma \alpha_\gamma^2 \right)^{1/2}$$

Elle est satisfaite si $\{u_\gamma\}$ est un système de **vaguelettes**

Décomposition en Ondelettes/Vaguelettes

Définition

Les fonctions continues $\{u_{j,k}\}$ forment un système de vaguelettes si $\exists \alpha > \beta > 0$ et $C < \infty$ t.q.

$$\begin{aligned} |u_{j,k}(x)| &\leq C2^{j/2}(1 + |2^j x - k|)^{-1-\alpha}), \\ \int u_{j,k}(x)dx &= 0, \\ |u_{j,k}(x)dx - u_{j,k}(x')| &\leq C2^{(\beta+1/2)j}|x - x'|^{\beta}. \end{aligned}$$

Décomposition en Ondelettes/Vaguelettes

- La théorie s'applique si K^* transforme un système d'ondelettes en un système de vaguelettes

Décomposition en Ondelettes/Vaguelettes

- La théorie s'applique si K^* transforme un système d'ondelettes en un système de vaguelettes
- C'est le cas si
$$KD_a \approx a^{\alpha} D_a K$$
où $D_a f(x) = f(ax)$ (dilatation), $\lambda_j = 2^{-j\alpha}$

Décomposition en Ondelettes/Vaguelettes

- La théorie s'applique si K^* transforme un système d'ondelettes en un système de vaguelettes
- C'est le cas si
$$KD_a \approx a^{\alpha} D_a K$$
  où $D_a f(x) = f(ax)$ (dilatation), $\lambda_j = 2^{-j\alpha}$
- Fonction d'échelle: $\phi_{j_1,l} = \sum_{j \leq j_1} \sum_k p_{(j_1,l),(j,k)} \psi_{j,k}$

Décomposition en Ondelettes/Vaguelettes

- La théorie s'applique si K^* transforme un système d'ondelettes en un système de vaguelettes
- C'est le cas si
$$KD_a \approx a^{\alpha} D_a K$$
où $D_a f(x) = f(ax)$ (dilatation), $\lambda_j = 2^{-j\alpha}$
- Fonction d'échelle: $\phi_{j_1,l} = \sum_{j \leq j_1} \sum_k p_{(j_1,l),(j,k)} \psi_{j,k}$
- On définit:
$$\delta_{j_1,l} = \sum_{j \leq j_1} \sum_k 2^{j\alpha} p_{(j_1,l),(j,k)} u_{j,k},$$
de sorte que $K^* \Delta_{j_1,l} = \phi_{j_1,l}$

Décomposition en Ondelettes/Vaguelettes

- On en déduit

$$
\begin{aligned}
||\sum_{j\leq j_1}\sum_{k} 2^{j\alpha} p_{(j_1,l),(j,k)} u_{j,k}||_{L_2} &\leq C2^{j_1\alpha}\left(\sum_{j\leq j_1}\sum_{k} p^2_{(j_1,l),(j,k)}\right)^{1/2} \\
&= C2^{j_1\alpha}
\end{aligned}
$$

Décomposition en Ondelettes/Vaguelettes

- On en déduit

$$\|\sum_{j\leq j_1}\sum_{k} 2^{j\alpha} p_{(j_1,l),(j,k)} u_{j,k}\|_{L_2} \leq C2^{j_1\alpha}\left(\sum_{j\leq j_1}\sum_{k} p^2_{(j_1,l),(j,k)}\right)^{1/2}$$

$$= C2^{j_1\alpha}$$

- On a la formule d'inversion inhomogène

$$f = \sum_{l}\langle Kf, \Delta_{j_1,l}\rangle \phi_{j_1,l} + \sum_{j\geq j_1,l} \lambda_j^{-1}\langle Kf, u_{j,k}\rangle \psi_{j,k}$$

- B.o.n. d'ondelettes: localisation temps-fréquence

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

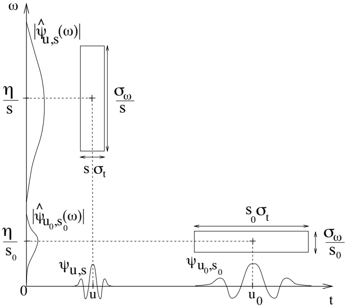


Figure 1.2: Time-frequency boxes of two wavelets $\psi_{u,s}$ and ψ_{u_0,s_0}. When the scale s decreases, the time support is reduced but the frequency spread increases and covers an interval that is shifted towards high frequencies.

Ondelettes et autres représentations

- Pavage du plan temps-fréquence

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

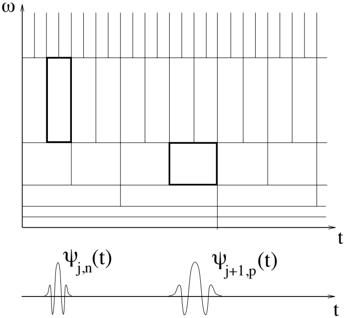


Figure 1.3: The time-frequency boxes of a wavelet basis define a tiling of the time-frequency plane.

Paquets d'ondelettes

- inhomogénéité/flexibilité fréquentielle

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

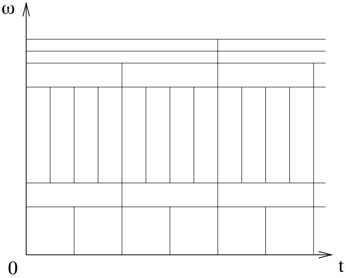


Figure 1.4: A wavelet packet basis divides the frequency axis in separate intervals of varying sizes. A tiling is obtained by translating in time the wavelet packets covering each frequency interval.

Paquets d'ondelettes

- Coifman, Meyer & Wickerhauser (1992)

Paquets d'ondelettes

- Coifman, Meyer & Wickerhauser (1992)
- Une structure d'arbre binaire

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

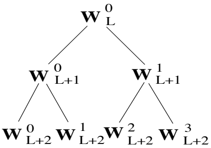


Figure 8.1: Binary tree of wavelet packet spaces.

Paquets d'ondelettes

- Coifman, Meyer & Wickerhauser (1992)
- Une structure d'arbre binaire

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

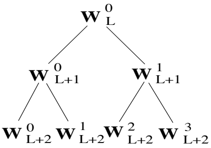


Figure 8.1: Binary tree of wavelet packet spaces.

- $\neq$ Analyse Multi-Résolution: $V_{j+1} = V_j \overset{\perp}{\bigoplus} W_j$

Paquets d'ondelettes

- Segmentation récursive de l'axe fréquentiel:
  $\Rightarrow$ partitions inhomogènes

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

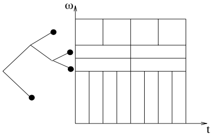


Figure 8.5: The wavelet packet tree on the left divides the frequency axis in several intervals. The Heisenberg boxes of the corresponding wavelet packet basis are on the right.

Paquets d'ondelettes

- Segmentation récursive de l'axe fréquentiel:
  $\Rightarrow$ partitions inhomogènes

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

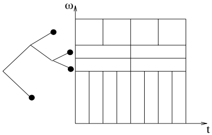


Figure 8.5: The wavelet packet tree on the left divides the frequency axis in several intervals. The Heisenberg boxes of the corresponding wavelet packet basis are on the right.

- $W_L^0 = \overset{\perp}{\bigoplus}_i W_{L_i}^{p_i}$

Paquets d'ondelettes

- Avec des arbres de profondeur $\leq J$, on peut construire: $\sim 2^{2^J}$ bases de $W_L^0 = V_L$!

Paquets d'ondelettes

- Avec des arbres de profondeur $\leq J$, on peut construire: $\sim 2^{2^J}$ bases de $W_L^0 = V_L$!
- Nécessité d'un algorithme pour choisir une "bonne" représentation

Paquets d'ondelettes

- Avec des arbres de profondeur $\leq J$, on peut construire: $\sim 2^{2^J}$ bases de $W_L^0 = V_L$!
- Nécessité d'un algorithme pour choisir une "bonne" représentation
- Adapté aux superpositions de chirps

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

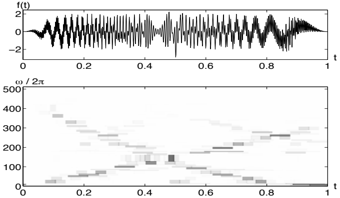


Figure 8.6: Wavelet packet decomposition of a multi-chirp signal. The darker the gray level of each Heisenberg box the larger the amplitude $|\langle f, \psi_j^p \rangle|$ of the corresponding wavelet packet coefficient.

Paquets de cosinus locaux

- Inhomogénéité/flexibilité temporelle

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

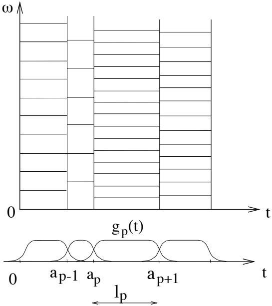
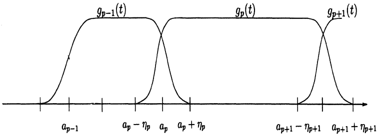


Figure 1.5: A local cosine basis divides the time axis with smooth windows $g_p(t)$. Multiplications with cosine functions translate these windows in frequency and yield a complete cover of the time-frequency plane.

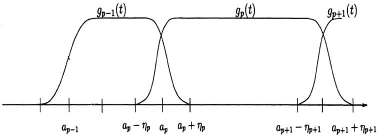
Paquets de cosinus locaux

- Partition inhomogène de l'axe temporel:



Paquets de cosinus locaux

- Partition inhomogène de l'axe temporel:



- Bases par blocs

Paquets de cosinus locaux

- Bases de cosinus IV:

Paquets de cosinus locaux

- Bases de cosinus IV:
  - Continu

$$\left\{\sqrt{2}\cos\left((k+\frac{1}{2})\pi t\right)\ ;\quad k\in\mathbb{N}\right\}\ \text{b.o.n. de } L_2([0,1])$$

Paquets de cosinus locaux

- Bases de cosinus IV:
  - Continu

$$\left\{ \sqrt{2}\cos\left((k+\frac{1}{2})\pi t\right)\ ; \quad k \in \mathbb{N} \right\} \text{ b.o.n. de } L_2([0,1])$$

  - Discret

$$\left\{ \sqrt{\frac{2}{N}}\cos\left(\frac{\pi}{N}(k+\frac{1}{2})\cdot(n+\frac{1}{2})\right)\ ; \quad 0 \leq k \leq N-1 \right\} \text{ b.o.n. de } \mathbb{R}^N$$

Paquets de cosinus locaux

- Bases de cosinus IV:
  - Continu

$$\left\{\sqrt{2}\cos\left((k+\frac{1}{2})\pi t\right)\ ;\ \ k\in\mathbb{N}\right\}\ \text{b.o.n. de } L_2([0,1])$$

  - Discret

$$\left\{\sqrt{\frac{2}{N}}\cos\left(\frac{\pi}{N}(k+\frac{1}{2})\cdot(n+\frac{1}{2})\right)\ ;\ \ 0\leq k\leq N-1\right\}\ \text{b.o.n. de } \mathbb{R}^N$$

- Algorithme de calcul rapide de la *DCT-IV*

Paquets de cosinus locaux

- **Challenge:** comment construire une b.o.n. de L_2 correspondant à une grille rectangulaire du plan temps-fréquence dont l'espacement varie avec le temps?

Paquets de cosinus locaux

- **Challenge:** comment construire une b.o.n. de L_2 correspondant à une grille rectangulaire du plan temps-fréquence dont l'espacement varie avec le temps?

Théorème de Balian-Low:

Il n'existe aucune b.o.n. de $L_2(\mathbb{R})$ de la forme

$$\left\{ h(t - nt_0) \exp(ik\omega_0); \;\; (n, k) \in \mathbb{Z}^2 \right\},$$

où h est une fenêtre différentiable.

Paquets de cosinus locaux

- **Challenge:** comment construire une b.o.n. de L_2 correspondant à une grille rectangulaire du plan temps-fréquence dont l'espacement varie avec le temps?

Théorème de Balian-Low:

Il n'existe aucune b.o.n. de $L_2(\mathbb{R})$ de la forme

$$\left\{h(t - nt_0)\exp(ik\omega_0); \;\; (n, k) \in \mathbb{Z}^2\right\},$$

où h est une fenêtre différentiable.

- **Malvar (1988):**
  Remplacer les exponentielles complexes par des cosinus IV.

Paquets de cosinus locaux

- La famille

$$\left\{ g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}} \cos\left(\pi(k+\frac{1}{2})\frac{t-a_p}{l_p}\right) \; ; \;\; (p,k) \in \mathbb{Z}\times\mathbb{N} \right\}$$

  est une b.o.n. de $L_2(\mathbb{R})$.

Paquets de cosinus locaux

- La famille

$$\left\{ g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}} \cos\left(\pi(k+\frac{1}{2})\frac{t-a_p}{l_p}\right) ;\ \ (p,k) \in \mathbb{Z} \times \mathbb{N} \right\}$$

  est une b.o.n. de $L_2(\mathbb{R})$.
- L'atome $g_{p,k}(t)$ est à support compact dans

$$[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$$

Paquets de cosinus locaux

- La famille

$$\left\{ g_{p,k}(t) = g_p(t)\sqrt{\frac{2}{l_p}}\cos\left(\pi(k+\frac{1}{2})\frac{t-a_p}{l_p}\right);\quad (p,k)\in\mathbb{Z}\times\mathbb{N}\right\}$$

  est une b.o.n. de $L_2(\mathbb{R})$.
- L'atome $g_{p,k}(t)$ est à support compact dans

$$[a_p - \eta_p, a_{p+1} + \eta_{p+1}]$$

- Sa transformée de Fourier est concentrée autour de

$$\omega_{p,k} = \pi(k+1/2)/l_p$$

  Elle est donnée par:

$$\widehat{g}_{p,k}(\omega) = \frac{e^{-ia_p\omega_{p,k}}}{2}\sqrt{\frac{2}{l_p}}\left(\widehat{g}_p(\omega-\omega_{p,k}) + \widehat{g}_p(\omega+\omega_{p,k})\right)$$

Paquets de cosinus locaux

- Pavage du plan TF par les rectangles de Heisenberg:

$$[a_p, a_{p+1}] \times [\omega_{p,k} - \frac{\pi}{2l_p}, \omega_{p,k} - \frac{\pi}{2l_p}]$$

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

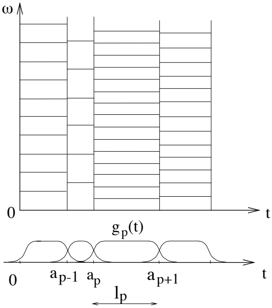


Figure 1.5: A local cosine basis divides the time axis with smooth windows $g_p(t)$. Multiplications with cosine functions translate these windows in frequency and yield a complete cover of the time-frequency plane.

Paquets de cosinus locaux - cas discret

- Soit $\{a_p\}_{p\in\mathbb{Z}}$ une suite de points de l'axe temporel t.q. $a_p + 1/2 \in \mathbb{Z}$,

$$\lim_{p\to+\infty} a_p = +\infty \text{ et } \lim_{p\to-\infty} a_p = -\infty.$$

Paquets de cosinus locaux - cas discret

- Soit $\{a_p\}_{p\in\mathbb{Z}}$ une suite de points de l'axe temporel t.q. $a_p + 1/2 \in \mathbb{Z}$,

$$\lim_{p\to+\infty} a_p = +\infty \text{ et } \lim_{p\to-\infty} a_p = -\infty.$$

- La famille

$$\left\{ g_{p,k}(n) = g_p(n)\sqrt{\frac{2}{l_p}}\cos\left(\pi(k+\frac{1}{2})\frac{n-a_p}{l_p}\right);\ \ p\in\mathbb{Z},\ 0\leq k<l_p \right\}$$

est une b.o.n. de $l_2(\mathbb{Z})$.

Paquets de cosinus locaux - structure arborescente

- Contraintes pratiques, algorithmiques:
  $\Rightarrow$ partitionnement **dyadique récursif** de l'axe temporel

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

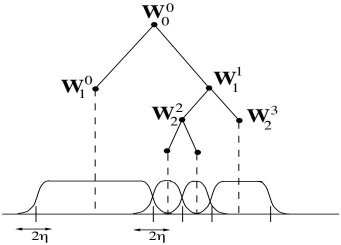


Figure 8.20: An admissible binary tree of local cosine spaces divides the time axis in windows of dyadic lengths.

Paquets de cosinus locaux et stationarité locale

- Un processus est **stationnaire** au second ordre ssi les exponentielles complexes diagonalisent sa matrice de covariance (*cf* convolution)

Paquets de cosinus locaux et stationarité locale

- Un processus est **stationnaire** au second ordre ssi les exponentielles complexes diagonalisent sa matrice de covariance (*cf* convolution)
- Un processus est **localement stationnaire** si il existe une base de Karhunen-Loeve formée de paquets de cosinus

Paquets de cosinus locaux et stationarité locale

- Un processus est **stationnaire** au second ordre ssi les exponentielles complexes diagonalisent sa matrice de covariance (*cf* convolution)
- Un processus est **localement stationnaire** si il existe une base de Karhunen-Loeve formée de paquets de cosinus
- Définition équivalente: $\forall (t,s) \in [u-\epsilon(u), u+\epsilon(u)]$,

$$\Gamma(t,s) = \gamma(\frac{t+s}{2}, t-s) \approx \gamma_u(t-s)$$

Paquets de cosinus locaux et stationarité locale

- Un processus est **stationnaire** au second ordre ssi les exponentielles complexes diagonalisent sa matrice de covariance (*cf* convolution)
- Un processus est **localement stationnaire** si il existe une base de Karhunen-Loeve formée de paquets de cosinus
- Définition équivalente: $\forall (t,s) \in [u-\epsilon(u), u+\epsilon(u)]$,

$$\Gamma(t,s) = \gamma(\frac{t+s}{2}, t-s) \approx \gamma_u(t-s)$$

- Comment rechercher une b.o.n. B t.q.

$$||B\Gamma B^* - diag(B\Gamma B^*)||_{HS}^2 = ||\Gamma||_{HS}^2 - ||diag(B\Gamma B^*)||_{HS}^2$$

soit minimum?

Choix d'une "meilleure" base/représentation

- Dictionnaire $\mathcal{D}$ de b.o.n. de $\mathcal{F} \sim \mathbb{R}^N$:

$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} \mathcal{B}_\lambda,$$

avec $\mathcal{B}_\lambda = \{g_{\lambda,m}\}_{1 \leq m \leq N}$

Choix d'une "meilleure" base/représentation

- Dictionnaire $\mathcal{D}$ de b.o.n. de $\mathcal{F} \sim \mathbb{R}^N$:

$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} \mathcal{B}_\lambda,$$

  avec $\mathcal{B}_\lambda = \{g_{\lambda,m}\}_{1 \leq m \leq N}$
- **Exemples:** arbres de paquets de cosinus locaux, d'ondelettes

Choix d'une "meilleure" base/représentation

- Dictionnaire $\mathcal{D}$ de b.o.n. de $\mathcal{F} \sim \mathbb{R}^N$:
$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} \mathcal{B}_\lambda,$$
avec $\mathcal{B}_\lambda = \{g_{\lambda,m}\}_{1 \leq m \leq N}$
- **Exemples:** arbres de paquets de cosinus locaux, d'ondelettes
- Meilleure base pour un signal $f \in \mathcal{F}$ donné? Choix **adaptatif**

Choix d'une "meilleure" base/représentation

- Dictionnaire $\mathcal{D}$ de b.o.n. de $\mathcal{F} \sim \mathbb{R}^N$:

$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} \mathcal{B}_\lambda,$$

avec $\mathcal{B}_\lambda = \{g_{\lambda,m}\}_{1 \leq m \leq N}$

- **Exemples:** arbres de paquets de cosinus locaux, d'ondelettes
- Meilleure base pour un signal $f \in \mathcal{F}$ donné? Choix **adaptatif**
- Approximation de f avec les $M \geq 1$ **meilleurs** vecteurs de $\mathcal{B}_\lambda$:

$$f_{\lambda,M} = \sum_{m \in I_{\lambda,m}} \langle f, g_{\lambda,m} \rangle g_{\lambda,m}$$

Choix d'une "meilleure" base/représentation

- Dictionnaire $\mathcal{D}$ de b.o.n. de $\mathcal{F} \sim \mathbb{R}^N$:
$$\mathcal{D} = \bigcup_{\lambda \in \Lambda} \mathcal{B}_\lambda,$$
avec $\mathcal{B}_\lambda = \{g_{\lambda,m}\}_{1 \leq m \leq N}$
- **Exemples:** arbres de paquets de cosinus locaux, d'ondelettes
- Meilleure base pour un signal $f \in \mathcal{F}$ donné? Choix **adaptatif**
- Approximation de f avec les $M \geq 1$ **meilleurs** vecteurs de $\mathcal{B}_\lambda$:
$$f_{\lambda,M} = \sum_{m \in I_{\lambda,m}} \langle f, g_{\lambda,m} \rangle g_{\lambda,m}$$
- L'erreur d'approximation quadratique s'écrit:
$$\begin{aligned} \epsilon_\lambda[M] &= \sum_{m \notin I_{\lambda,m}} \langle f, g_{\lambda,m} \rangle^2 \\ &= ||f||^2 - \sum_{m \in I_{\lambda,m}} \langle f, g_{\lambda,m} \rangle^2 \end{aligned}$$

Choix d'une "meilleure" base/représentation

- Un **ordre partiel** sur $\mathcal{D}$

Définition

Soit $(\lambda_1, \lambda_2) \in \Lambda^2$. La base $\mathcal{B}_{\lambda_1}$ est meilleure que la base $\mathcal{B}_{\lambda_2}$ pour représenter le signal f ssi $\forall M \geq 1$:

$$\epsilon_{\lambda_1}[M] \leq \epsilon_{\lambda_2}[M]$$

Choix d'une "meilleure" base/représentation

- Un **ordre partiel** sur $\mathcal{D}$

Définition

Soit $(\lambda_1, \lambda_2) \in \Lambda^2$. La base $\mathcal{B}_{\lambda_1}$ est meilleure que la base $\mathcal{B}_{\lambda_2}$ pour représenter le signal f ssi $\forall M \geq 1$:

$$\epsilon_{\lambda_1}[M] \leq \epsilon_{\lambda_2}[M]$$

- **Caractérisation**

Proposition

Soit $(\lambda_1, \lambda_2) \in \Lambda^2$. La base $\mathcal{B}_{\lambda_1}$ est meilleure que la base $\mathcal{B}_{\lambda_2}$ pour représenter le signal f ssi on a pour toute fonction concave Φ:

$$\sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda_1,m}\rangle^2}{||f||^2}\right) \leq \sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda_2,m}\rangle^2}{||f||^2}\right).$$

Choix d'une "meilleure" base/représentation

- En général, on fixe une fonction concave et on cherche à minimiser le **coût** de l'approximation de f dans une base candidate $\mathcal{B}_\lambda$:

$$C(f, \mathcal{B}_\lambda) = \sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda,m}\rangle^2}{||f||^2}\right)$$

Choix d'une "meilleure" base/représentation

- En général, on fixe une fonction concave et on cherche à minimiser le **coût** de l'approximation de f dans une base candidate $\mathcal{B}_\lambda$:

$$C(f, \mathcal{B}_\lambda) = \sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda,m}\rangle^2}{||f||^2}\right)$$

- Meilleure base au sens de la fonction de coût choisie pour f:

$$\mathcal{B}_{\lambda^*} = \arg\min_{\mathcal{B}_\lambda} C(f, \mathcal{B}_\lambda)$$

Choix d'une "meilleure" base/représentation

- En général, on fixe une fonction concave et on cherche à minimiser le **coût** de l'approximation de f dans une base candidate $\mathcal{B}_\lambda$:

$$C(f, \mathcal{B}_\lambda) = \sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda,m}\rangle^2}{||f||^2}\right)$$

- Meilleure base au sens de la fonction de coût choisie pour f:

$$\mathcal{B}_{\lambda^*} = \arg\min_{\mathcal{B}_\lambda} C(f, \mathcal{B}_\lambda)$$

- Exemples:
  - Entropie: $\Phi(x) = -x\log(x)$

Choix d'une "meilleure" base/représentation

- En général, on fixe une fonction concave et on cherche à minimiser le **coût** de l'approximation de f dans une base candidate $\mathcal{B}_\lambda$:

$$C(f, \mathcal{B}_\lambda) = \sum_{m=1}^{N} \Phi\left(\frac{\langle f, g_{\lambda,m}\rangle^2}{||f||^2}\right)$$

- Meilleure base au sens de la fonction de coût choisie pour f:

$$\mathcal{B}_{\lambda^*} = \arg\min_{\mathcal{B}_\lambda} C(f, \mathcal{B}_\lambda)$$

- Exemples:
    - Entropie: $\Phi(x) = -x\log(x)$
    - coût l^p avec $p \leq 2$: $\Phi(x) = x^{p/2}$

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux
  - Pour un signal de dimension N: $\#\mathcal{D} \sim 2^{N/2}$

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux
  - Pour un signal de dimension N: $\#\mathcal{D} \sim 2^{N/2}$
  - Recherche **exhaustive**: $O(N2^{N/2})$ opérations!

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux
  - Pour un signal de dimension N: $\#\mathcal{D} \sim 2^{N/2}$
  - Recherche **exhaustive**: $O(N2^{N/2})$ opérations!
- Pour des dictionnaires à structure arborescente: stratégie **"bottom-up"** de Coifman & Wickerhauser (1996):

$$O(N\log_2(N))) \text{ opérations seulement!}$$

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux
  - Pour un signal de dimension N: $\#\mathcal{D} \sim 2^{N/2}$
  - Recherche **exhaustive**: $O(N2^{N/2})$ opérations!
- Pour des dictionnaires à structure arborescente: stratégie **"bottom-up"** de Coifman & Wickerhauser (1996):

$$O(N\log_2(N)))\text{ opérations seulement!}$$

- **Structure arborescente:** en plus de la base $\mathcal{B}_{j,k}$ de $W_{j,k} = W_{j+1,2k} \oplus W_{j+1,2k+1}$, on peut en construire une en réunissant les bases des sous-espaces correspondant aux descendants:

$$\mathcal{B}_{j+1,2k} \bigcup \mathcal{B}_{j+1,2k+1}$$

Algorithmes de recherche d'une "meilleure" base

- **Exemple:** dictionnaire $\mathcal{D}$ des bases de paquets de cosinus locaux
  - Pour un signal de dimension N: $\#\mathcal{D} \sim 2^{N/2}$
  - Recherche **exhaustive**: $O(N2^{N/2})$ opérations!
- Pour des dictionnaires à structure arborescente: stratégie **"bottom-up"** de Coifman & Wickerhauser (1996):

$$O(N\log_2(N))) \text{ opérations seulement!}$$

- **Structure arborescente:** en plus de la base $\mathcal{B}_{j,k}$ de $W_{j,k} = W_{j+1,2k} \oplus W_{j+1,2k+1}$, on peut en construire une en réunissant les bases des sous-espaces correspondant aux descendants:

$$\mathcal{B}_{j+1,2k} \bigcup \mathcal{B}_{j+1,2k+1}$$

- **Additivité:**

$$C(f, \mathcal{B}_{j+1,2k} \bigcup \mathcal{B}_{j+1,2k+1}) = C(f, \mathcal{B}_{j+1,2k}) + C(f, \mathcal{B}_{j+1,2k+1})$$

Coifman & Wickerhauser (1996)

- **"Tournoi"**

Proposition

Soit $\mathcal{MB}_{j,k}$ la meilleure base de l'espace $W_{j,k}$. On a

$$\mathcal{MB}_{j,k} = \mathcal{B}_{j,k} \text{ si } C(f, \mathcal{MB}_{j+1,2k}) + C(f, \mathcal{MB}_{j+1,2k+1}) \geq C(f, \mathcal{B}_{j+1,2k})$$

et

$$\mathcal{MB}_{j,k} = \mathcal{MB}_{j+1,2k} \bigcup \mathcal{MB}_{j+1,2k+1} \text{ sinon }.$$

Coifman & Wickerhauser (1996)

- **"Tournoi"**

Proposition

Soit $\mathcal{MB}_{j,k}$ la meilleure base de l'espace $W_{j,k}$. On a

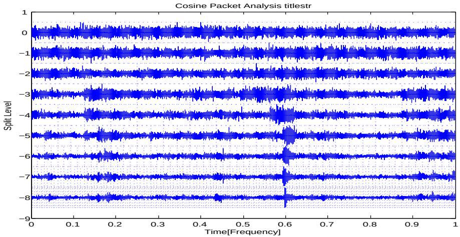
$$\mathcal{MB}_{j,k} = \mathcal{B}_{j,k} \text{ si } C(f, \mathcal{MB}_{j+1,2k}) + C(f, \mathcal{MB}_{j+1,2k+1}) \geq C(f, \mathcal{B}_{j+1,2k})$$

et

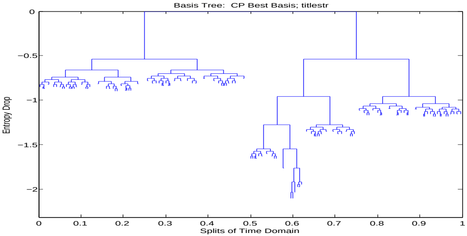
$$\mathcal{MB}_{j,k} = \mathcal{MB}_{j+1,2k} \bigcup \mathcal{MB}_{j+1,2k+1} \text{ sinon .}$$

- **Principe:** On remonte l'arbre de bas en haut en effectuant les tournois le long des branches.

Coifman & Wickerhauser (1996)



Coifman & Wickerhauser (1996)



Application: estimation de la covariance d'un processus localement stationnaire

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

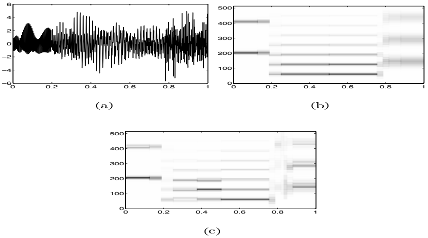


Figure 10.19: (a): One realization of a process X_N that is stationary on $[0, 0.2]$, $[0.2, 0.78]$ and $[0.78, 1]$, with $N = 1024$. (b): Heisenberg boxes of the best local cosine basis computed with $L = 500$ realizations of this locally stationary process. Grey levels are proportional to the estimated spectrum. (c): Best local cosine basis calculated with $L = 3$ realizations.

Approximation par poursuite

- **Complexité** du signal à approcher $\Rightarrow$ **flexibilité** dans le choix des atomes "temps-fréquence" utilisés dans la représentation

Approximation par poursuite

- **Complexité** du signal à approcher $\Rightarrow$ **flexibilité** dans le choix des atomes "temps-fréquence" utilisés dans la représentation
- **Idée:** abandonner la contrainte d'orthogonalité

Approximation par poursuite

- **Complexité** du signal à approcher $\Rightarrow$ **flexibilité** dans le choix des atomes "temps-fréquence" utilisés dans la représentation
- **Idée:** abandonner la contrainte d'orthogonalité
- Constituer des bases non nécéssairement o.n. à partir de la collection de tous les vecteurs du dictionnaire $\mathcal{D} = \{g_p\}_{1 \leq p \leq P}$, $P >> N$

Approximation par poursuite

- **Complexité** du signal à approcher $\Rightarrow$ **flexibilité** dans le choix des atomes "temps-fréquence" utilisés dans la représentation
- **Idée:** abandonner la contrainte d'orthogonalité
- Constituer des bases non nécéssairement o.n. à partir de la collection de tous les vecteurs du dictionnaire $\mathcal{D} = \{g_p\}_{1 \leq p \leq P}$, $P >> N$
- Approximation du signal f avec M vecteurs:

$$f_M = \sum_{m=1}^{M} \alpha_{p_m} g_{p_m}$$

Poursuite de base - Chen & Donoho (1996)

- Soit $\mathcal{B} = \{g_{p_m}\}_{1 \leq m \leq N}$ une base de vecteurs du dictionnaire:

$$f = \sum_{m=1}^{N} \alpha_{p_m} g_{p_m}$$

Poursuite de base - Chen & Donoho (1996)

- Soit $\mathcal{B} = \{g_{p_m}\}_{1 \leq m \leq N}$ une base de vecteurs du dictionnaire:

$$f = \sum_{m=1}^{N} \alpha_{p_m} g_{p_m}$$

- On détermine la base $\mathcal{B}$ minimisant le coût l_1:

$$C(f, \mathcal{B}) = \frac{\sum_{m=1}^{N} |\alpha_{p_m}|}{||f||}$$

Poursuite de base - Chen & Donoho (1996)

- Soit $\mathcal{B} = \{g_{p_m}\}_{1 \leq m \leq N}$ une base de vecteurs du dictionnaire:

$$f = \sum_{m=1}^{N} \alpha_{p_m} g_{p_m}$$

- On détermine la base $\mathcal{B}$ minimisant le coût l_1:

$$C(f, \mathcal{B}) = \frac{\sum_{m=1}^{N} |\alpha_{p_m}|}{||f||}$$

- Norme l_1 et **sparsité**

Poursuite de base - Chen & Donoho (1996)

- Soit $\mathcal{B} = \{g_{p_m}\}_{1 \leq m \leq N}$ une base de vecteurs du dictionnaire:

$$f = \sum_{m=1}^{N} \alpha_{p_m} g_{p_m}$$

- On détermine la base $\mathcal{B}$ minimisant le coût l_1:

$$C(f, \mathcal{B}) = \frac{\sum_{m=1}^{N} |\alpha_{p_m}|}{||f||}$$

- Norme l_1 et **sparsité**
- Un programme linéaire pour le calcul de la solution du problème non-linéaire:

$$\arg\min_{\mathcal{B}} C(f, \mathcal{B})$$

Poursuite de base - Chen & Donoho (1996)

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

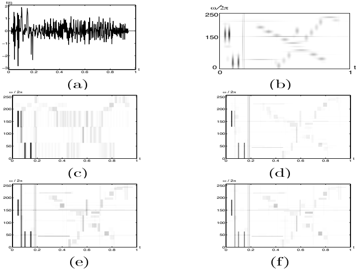


Figure 9.11: (a): Signal synthesized with a sum of chirps, truncated sinusoids, short time transients and Diracs. The time-frequency images display the atoms selected by different adaptive time-frequency transforms. The darkness is proportional to the coefficient amplitude. (b): Gabor matching pursuit. Each dark blob is the Wigner-Ville distribution of a selected Gabor atom. (c): Heisenberg boxes of a best wavelet packet basis calculated with Daubechies 8 filter. (d): Wavelet packet basis pursuit. (e): Wavelet packet matching pursuit. (f): Wavelet packet orthogonal matching pursuit.

Matching Pursuit - Mallat & Zhong (1996)

- Le dictionnaire des paquets de cosinus locaux contient $P = N\log_2(N)$ vecteurs

Matching Pursuit - Mallat & Zhong (1996)

- Le dictionnaire des paquets de cosinus locaux contient $P = N \log_2(N)$ vecteurs
- Coût **prohibitif** de la poursuite de base: programmation linéaire sur les P vecteurs de $\mathcal{D}$

Matching Pursuit - Mallat & Zhong (1996)

- Le dictionnaire des paquets de cosinus locaux contient $P = N\log_2(N)$ vecteurs
- Coût **prohibitif** de la poursuite de base: programmation linéaire sur les P vecteurs de $\mathcal{D}$
- **Poursuite par projection** en Statistique:

Matching Pursuit - Mallat & Zhong (1996)

- Le dictionnaire des paquets de cosinus locaux contient $P = N\log_2(N)$ vecteurs
- Coût **prohibitif** de la poursuite de base: programmation linéaire sur les P vecteurs de $\mathcal{D}$
- **Poursuite par projection** en Statistique:
  1. **Initialization:** On choisit de projeter f sur g_{γ_0} dans $\mathcal{D}$,

     $$f = \langle f, g_{\gamma_0}\rangle g_{\gamma_0} + Rf,$$

     de façon à minimiser:

     $$||Rf||^2 = ||f||^2 - |\langle f, g_{\gamma_0}\rangle|^2$$

     ou, de manière équivalente, afin de maximiser:

     $$|\langle f, g_{\gamma_0}\rangle|^2$$

Matching Pursuit - Mallat & Zhong (1996)

- Le dictionnaire des paquets de cosinus locaux contient $P = N\log_2(N)$ vecteurs
- Coût **prohibitif** de la poursuite de base: programmation linéaire sur les P vecteurs de $\mathcal{D}$
- **Poursuite par projection** en Statistique:
  1. **Initialization:** On choisit de projeter f sur g_{γ_0} dans $\mathcal{D}$,

     $$f = \langle f, g_{\gamma_0} \rangle g_{\gamma_0} + Rf,$$

     de façon à minimiser:

     $$||Rf||^2 = ||f||^2 - |\langle f, g_{\gamma_0} \rangle|^2$$

     ou, de manière équivalente, afin de maximiser:

     $$|\langle f, g_{\gamma_0} \rangle|^2$$

  2. **Poursuite:** actualiser $f \to Rf$, appliquer la règle précédente au résidu

Matching Pursuit

Proposition

L'algorithme converge à une vitesse exponentielle: $\exists \rho \in]0,1[$ t.q. $\forall m \geq 1$

$$||R_m f|| \leq \rho^m ||f||$$

Matching Pursuit

Proposition

L'algorithme converge à une vitesse exponentielle: $\exists \rho \in]0,1[$ t.q. $\forall m \geq 1$

$$||R_m f|| \leq \rho^m ||f||$$

- Règle d'arrêt: si $||R_m f||^2 \leq \epsilon ||f||^2$
- Mise à jour:

$$\langle R_{m+1} f, g_\gamma \rangle = \langle R_m f, g_\gamma \rangle - \langle R_m f, g_{\gamma_m} \rangle \cdot \langle g_{\gamma_m}, g_\gamma \rangle$$

Matching Pursuit

Proposition

L'algorithme converge à une vitesse exponentielle: $\exists \rho \in]0,1[$ t.q. $\forall m \geq 1$

$$||R_m f|| \leq \rho^m ||f||$$

- Règle d'arrêt: si $||R_m f||^2 \leq \epsilon ||f||^2$
- Mise à jour:

$$\langle R_{m+1} f, g_\gamma \rangle = \langle R_m f, g_\gamma \rangle - \langle R_m f, g_{\gamma_m} \rangle \cdot \langle g_{\gamma_m}, g_\gamma \rangle$$

- Rapide si beaucoup de coefficients $\langle g_{\gamma_m}, g_\gamma \rangle$ sont nuls

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

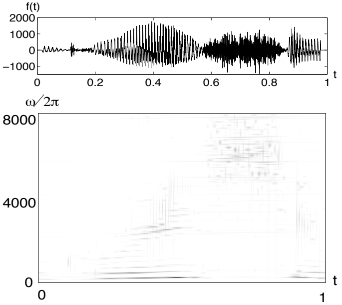


Figure 9.12: Speech recording of the word "greasy" sampled at $16kHz$. In the time-frequency image, the dark blobs of various sizes are the Wigner-Ville distributions of a Gabor functions selected by the matching pursuit.

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"
- Ingrédient commun: dictionnaire $\mathcal{D}$ ($\forall g \in \mathcal{D},\ ||g|| = 1$)

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"
- Ingrédient commun: dictionnaire $\mathcal{D}$ ($\forall g \in \mathcal{D}$, $||g|| = 1$)
- Algorithmes **récursifs**: $f_0 = 0$, à l'étape k, l'approximation f_k est obtenue à partir de f_{k_1} et son résidu $r_{k-1} = f - f_{k-1}$.

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"
- Ingrédient commun: dictionnaire $\mathcal{D}$ ($\forall g \in \mathcal{D}$, $||g|| = 1$)
- Algorithmes **récursifs**: $f_0 = 0$, à l'étape k, l'approximation f_k est obtenue à partir de f_{k_1} et son résidu $r_{k-1} = f - f_{k-1}$.
- PGA et OGA: à chaque pas, on détermine

$$g_k = \arg\max_{g \in \mathcal{D}} |\langle r_{k-1}, g \rangle|,$$

  et on pose alors

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"
- Ingrédient commun: dictionnaire $\mathcal{D}$ ($\forall g \in \mathcal{D},\ ||g|| = 1$)
- Algorithmes **récursifs**: $f_0 = 0$, à l'étape k, l'approximation f_k est obtenue à partir de f_{k_1} et son résidu $r_{k-1} = f - f_{k-1}$.
- PGA et OGA: à chaque pas, on détermine

$$g_k = \arg\max_{g \in \mathcal{D}} |\langle r_{k-1}, g \rangle|,$$

  et on pose alors
  - $f_k = f_{k-1} + \langle r_{k-1}, g_k \rangle g_k$ dans le cas PGA

Algorithmes "gloutons" - nomenclature

- Quatre types d'algorithmes "gloutons"
  1. PGA "glouton pur"
  2. OGA "glouton orthogonal"
  3. RGA "glouton relaxé"
  4. SPA "projection pas à pas"
- Ingrédient commun: dictionnaire $\mathcal{D}$ ($\forall g \in \mathcal{D}$, $||g|| = 1$)
- Algorithmes **récursifs**: $f_0 = 0$, à l'étape k, l'approximation f_k est obtenue à partir de f_{k_1} et son résidu $r_{k-1} = f - f_{k-1}$.
- PGA et OGA: à chaque pas, on détermine

$$g_k = \arg\max_{g \in \mathcal{D}} |\langle r_{k-1}, g \rangle|,$$

  et on pose alors
  - $f_k = f_{k-1} + \langle r_{k-1}, g_k \rangle g_k$ dans le cas PGA
  - $f_k = proj_{V_k} f$ avec $V_k = g_1, \ldots, g_k$ dans le cas OGA

Algorithmes "gloutons" - nomenclature

- Dans le cas RGA, l'approximation à l'étape k a la forme

$$f_k = \alpha_k f_{k-1} + \beta_k g_k,$$

avec $(\alpha_k, \beta_k, g_k) = \arg\min_{(\alpha,\beta,g)} ||f - \alpha f_{k-1} - \beta g||$.

Algorithmes "gloutons" - nomenclature

- Dans le cas RGA, l'approximation à l'étape k a la forme

$$f_k = \alpha_k f_{k-1} + \beta_k g_k,$$

avec $(\alpha_k, \beta_k, g_k) = \arg\min_{(\alpha,\beta,g)} ||f - \alpha f_{k-1} - \beta g||$.

- Dans le cas SPA, on recherche $g \in \mathcal{D}$ de façon à minimiser

$$||f - proj_{Vect(g_1,\ldots,g_{k-1},g)} f||$$

Algorithmes "gloutons" - nomenclature

- Dans le cas RGA, l'approximation à l'étape k a la forme

$$f_k = \alpha_k f_{k-1} + \beta_k g_k,$$

avec $(\alpha_k, \beta_k, g_k) = \arg\min_{(\alpha,\beta,g)} ||f - \alpha f_{k-1} - \beta g||$.

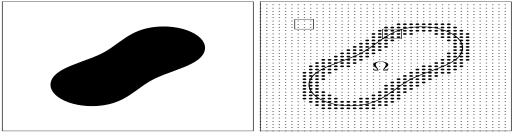
- Dans le cas SPA, on recherche $g \in \mathcal{D}$ de façon à minimiser

$$||f - proj_{Vect(g_1,\ldots,g_{k-1},g)} f||$$

- En pratique, pour OGA et SPA, les projections orthogonales successives sont calculées par Gramm-Schmidt.

- Elaborer des procédures permettant de décrire les singularités pour le débruitage, la compression, *etc.*
- Pour un signal unidimensionnel: ondelettes, paquets d'ondelettes, *etc.*
- Pour une image?

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)



(a) (b)

Analyse harmonique algorithmique

A Wavelet Tour of Signal Processing
Stéphane Mallat, Academic Press 1999 (2nd edition)

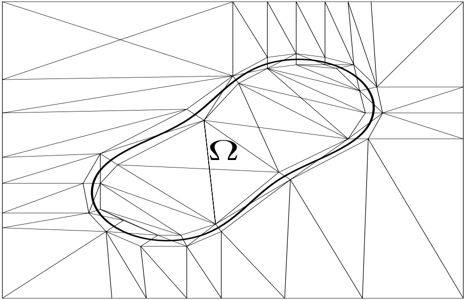


Figure 9.6: A piecewise linear approximation of $f = \mathbf{1}_\Omega$ is optimized with a triangulation whose triangles are narrow in the direction where f is discontinuous, along the border $\partial\Omega$.

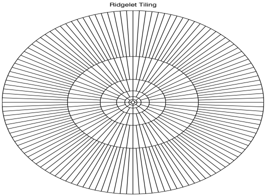
Analyse harmonique algorithmique

- Exemples: fonction $f(x_1, x_2)$ ayant des singularités le long de droites

Analyse harmonique algorithmique

- Exemples: fonction $f(x_1, x_2)$ ayant des singularités le long de droites
- Développement dans une base de *ridgelets*:

$$\psi_{a,b,\theta}(x) = a^{-1/2}\psi(a^{-1}(\langle x, (\cos\theta, \sin\theta)\rangle - b))$$



- A chaque type de singularité, sa classe de représentation: Curvelets, wedgelets, bandelets, *etc.*