

Sujet M2 2025/2026: estimation adaptative et optimale par méthodes d'ondelettes

Mathieu Sart
Institut Camille Jordan
mathieu.sart@univ-st-etienne.fr

Les bases d'ondelettes sont des bases particulières de $\mathbb{L}^2(\mathbb{R})$ obtenues en faisant des translations et des dilatations de deux fonctions spécifiques ϕ et ψ . Elles donnent naissance à la base $\{\phi_k, k \in \mathbb{Z}\} \cup \{\psi_{j,k}, j \geq 0, k \in \mathbb{Z}\}$ où pour tout $x \in \mathbb{R}$,

$$\begin{aligned}\phi_k(x) &= \phi(x - k) \\ \psi_{j,k}(x) &= 2^{j/2} \psi(2^j x - k).\end{aligned}$$

Ainsi, toute fonction f de carré intégrable admet la représentation

$$f = \sum_{k \in \mathbb{Z}} \alpha_k \phi_k + \sum_{j=0}^{\infty} \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}$$

où les α_k et les $\beta_{j,k}$ sont des réels.

L'intérêt des ondelettes est qu'elles permettent une représentation parcimonieuse de la fonction d'intérêt. On peut ainsi approcher convenablement une fonction f avec un petit nombre de coefficients. Mathématiquement, toute fonction f "régulière" pourra être très bien approchée par une fonction \bar{f} de la forme

$$\bar{f} = \sum_{k \in K} \alpha_k \phi_k + \sum_{j=0}^J \sum_{k \in K_j} \beta_{j,k} \psi_{j,k} \quad (1)$$

où $J \geq 0$, et où K et les K_j sont des ensembles d'entiers de petite taille.

En statistiques non-paramétriques, cette propriété de compression peut être exploitée pour estimer une fonction f d'intérêt sous des hypothèses faibles. L'idée est de partir de la représentation (1) dans laquelle on va remplacer les coefficients α_k et $\beta_{j,k}$ inconnus par des versions estimées. Cela conduit à l'estimateur

$$\hat{f} = \sum_{k \in K} \hat{\alpha}_k \phi_k + \sum_{j=0}^J \sum_{k \in K_j} \hat{\beta}_{j,k} \psi_{j,k}$$

de f . Si l'approximation (1) est d'autant meilleure que les ensembles K et K_j sont gros, il n'en est pas de même pour le risque de l'estimateur \hat{f} . Le choix de ces ensembles doit être issu d'un compromis biais-variance.

Réaliser ce compromis de manière optimale et ainsi avoir les vitesses minimax de convergence peut cependant être difficile dans certains modèles comme celui de la densité. L'objectif de ce stage est de se familiariser avec les ondelettes, les méthodes classiques de la littérature pour réaliser le compromis, ainsi que celles plus récentes comme celle de [Sar25].

Ce sujet s'adresse en priorité aux étudiants souhaitant poursuivre en thèse. Un projet de thèse financé sur la théorie minimax pour les contraintes de forme et de régularité commençant en septembre 2026 devrait probablement être ouvert prochainement. Toutes les candidatures ayant un solide background en mathématiques seront examinées, y compris celles n'ayant pas suivi de cours approfondis de statistiques mathématiques.

References

- [Sar25] Mathieu Sart. Non-linear wavelet density estimation on the real line. *Bernoulli*, 2025. To appear.