

# Approximation de faible rang de la matrice de Gram et détection de ruptures sur des signaux de grande taille

## Projet

Dans le contexte du machine/statistical learning, les approches basées sur les noyaux semi-définis positifs (noyaux reproduisants) sont très répandues.

Elles présentent entre autres l'avantage de pouvoir traiter des types variés de données (graphes, séquences d'ADN, ...).

Une limitation de ces approches basées sur les noyaux est qu'elles induisent un coût algorithmique souvent trop important lié aux calcul et stockage de la matrice de Gram de taille  $n \times n$ .

Les objectifs du stage sont :

1. mettre en œuvre et analyser plusieurs techniques d'approximation de la matrice de Gram par une matrice de faible rang (de type Nystrom par exemple),
2. développer une nouvelle approche permettant de choisir les colonnes de la matrice de Gram utilisées pour en bâtir une approximation de faible rang,
3. analyser et mettre en oeuvre l'approche globale sur des exemples.

## Prérequis

Outre de bonnes compétences en probabilités et statistique, le candidat utilisera R ou matlab pour réaliser des simulations afin de vérifier empiriquement ses résultats.

*Encadrant* : Alain Celisse, maître de conférences.

*Length*: 4-6 mois.

*Opportunity*: Le stage pourrait donner lieu à une thèse.

*Laboratory*: MODAL équipe-projet Inria, Lille.

*Contact*: Alain Celisse ([celisse@math.univ-lille1.fr](mailto:celisse@math.univ-lille1.fr)).

*Suite*: Le sujet proposé pourrait faire l'objet d'une thèse.