
Analyse statistique de données à structure de graphes, via des processus de diffusion de la chaleur.

Etienne Lasalle*¹

¹Laboratoire de Mathématiques d'Orsay – Université Paris-Saclay, Centre National de la Recherche Scientifique : UMR8628 – France

Résumé

Lorsqu'on travaille avec des graphes pondérés, on peut interpréter les poids des arêtes comme des conductivités thermiques: la chaleur se diffuse plus rapidement le long des arêtes ayant un poids important. Dès lors, on peut comparer les graphes en comparant leur répartition de chaleur au bout d'un temps de diffusion t . Ce paramètre t , jouant le rôle d'un paramètre d'échelle, doit être minutieusement choisi pour s'assurer des comparaisons pertinentes.

A l'opposé de certains travaux considérant un temps de diffusion fixé arbitrairement ou choisi à partir des données, on se propose de prendre en compte tout le processus de diffusion de la chaleur. Pour cela, on définit des processus à valeurs réelles indexés par tous les temps de diffusion dans $[0, T]$, concaténant toutes les comparaisons faites aux différentes échelles. Dans cet exposé, je présenterai ces processus de comparaisons, ainsi que leurs propriétés statistiques. Notamment, on montre qu'ils vérifient un théorème central limite fonctionnel et qu'ils admettent des approximations gaussiennes. En pratique, ces résultats permettent de construire des bandes de confiance asymptotiques ainsi que des tests à deux échantillons.

*Intervenant