

Éditorial du numéro spécial sur l'analyse de mélanges

Title: Editorial of the Special Issue on Mixture Analysis

Gilles Celeux¹

Les modèles de mélange jouent un rôle important en inférence statistique depuis de nombreuses années. Ces modèles à structure cachée permettent la prise en compte d'hétérogénéité, une situation courante dans beaucoup de données. Ainsi, les modèles de mélange peuvent être vus comme des outils d'estimation semi-paramétriques de densité de probabilité et sont par ailleurs l'outil de référence pour les modèles de classification. Ils trouvent des applications dans des contextes variés comme la classification non supervisée de données indépendantes, les modèles de Markov cachés, la classification de réseaux, les modèles de mélange d'experts, la classification croisée... Ils induisent des questions importantes, difficiles et intéressantes, dont le choix d'un nombre pertinent de composants n'est pas la moindre.

De fait, les modèles de mélange ont été l'objet de livres de référence au long des années. Le livre de [Titterton et al., 1985] peut être recommandé notamment pour sa présentation des bases mathématiques du modèle de mélange. Le livre de [McLachlan and Peel, 2000] peut être recommandé notamment pour sa présentation du modèle de mélange comme outil de classification non supervisée. Le livre de [Frühwirth-Schnatter, 2006] peut être recommandé notamment pour son orientation vers l'inférence bayésienne. L'activité éditoriale sur les modèles de mélange reste importante. On peut ainsi citer le *Handbook of Mixture Analysis* de [Frühwirth-Schnatter et al., 2018] qui présente un large tour d'horizon des méthodes et des applications du modèle de mélange. et la monographie à paraître prochainement sur la classification par le modèle de mélange [Bouveyron et al., 2019].

Ce numéro spécial traite d'aspects variés de cet important champ de recherche : la classification de réseaux dynamiques, le modèle de mélange d'experts, les tests sur le nombre de composants d'un mélange gaussien univarié, des procédures bootstrap pour évaluer le nombre de composants d'un mélange et une application en génomique. Il est à noter que tous les articles de ce numéro spécial considèrent le problème de sélection de modèles et en particulier le problème crucial du choix du nombre de composants d'un mélange.

Ricardo Rastelli s'intéresse à la maximisation directe de la vraisemblance intégrée exacte dans un contexte non informatif pour le *Stochastic Block Model* appliqué à des réseaux dynamiques.

Faïcel Chamroukhi et Bao-Tuyen Hyung proposent des procédures d'estimation et de sélection de variables régularisées pour des mélanges d'experts afin d'obtenir des solutions parcimonieuses.

¹ Inria Saclay-Île-de-France, Orsay, F-91405, France

Didier Chauveau, Bernard Garel et Sophie Mercier reviennent sur les problème classique de tester le nombre de composants d'un mélange gaussien univarié d'un point de vue pratique.

Zhivko Taushanov et André Berchtold proposent différentes procédures bootstrap pour sélectionner un modèle de mélange. Ils concentrent leurs illustrations sur le modèle *Hidden Mixture Transition Distribution* pour la classification de données longitudinales.

Christine Keribin, Yi Lia, Tatiana Popova et Yves Rozenholc présentent une application du modèle de mélanlge à des données génétiques. Ils se concentrent sur la sélection d'un modèle pertinent et montre l'intérêt de l'heuristique de pente dans le contexte de leur étude.

Je remercie les auteurs pour leur contributions et j'espère que les lecteurs du Journal de la SFdS seront intéressés par ce numéro spécial.

Références

- [Bouveyron et al., 2019] Bouveyron, C., Celeux, G., Murphy, T. B., and Raftery, A. E. (2019). *Model-based Clustering and Classification for Data Science*. Cambridge University Press. (Forthcoming).
- [Frühwirth-Schnatter, 2006] Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer-Verlag.
- [Frühwirth-Schnatter et al., 2018] Frühwirth-Schnatter, S., Celeux, G., and Robert, C. P. (2018). *Handbook of Mixture Analysis*. CRC Press.
- [McLachlan and Peel, 2000] McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. John Wiley.
- [Titterton et al., 1985] Titterton, D. M., Smith, A. F. M., and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture distributions*. John Wiley.