Copula-based models for the inference of multi-omics regulation networks

Ekaterina Tomilina*[1,2]

[1]Mathématiques et Informatique Appliquées du Génome à l'Environnement [Jouy-En-Josas] – Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France
[2]Génétique Animale et Biologie Intégrative – Université Paris-Saclay, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France

Résumé

When inferring multi-omics regulation networks, one can be quickly faced with statistical challenges arising from the heterogeneous and high-dimensional nature of the data. In that case, a good strategy can be to use a Gaussian copula model in order to identify independent modules within the network. When the marginal laws of the variables are unknown, a semi-parametric estimation of the correlation matrix of the copula can be done via the pairwise maximum likelihood with respect to a mix of Lebesgue and counting measures, leading to asymptotically normal estimators. We show that a block structure of the correlation matrix corresponds to the presence of independent modules within the network. We discuss thresholding methods in order to control the false discovery rate of non-zero coefficients. We illustrate the benefits of copula-based models with respect to standard approaches on simulations and real data.

---

*Intervenant