Multiple testing of paired null hypotheses using a latent graph model

Tabea Rebafka[1] **and Etienne Roquain**[1] **and Fanny Villers**[1]

[1] *Sorbonne Université, Paris, France, fanny.villers@upmc.fr*

We consider pairwise measures between n entities of interest and we want to decide which pairs of entities are significantly related. This multiple testing problem can also be seen as follows : interactions between entities are described via noisy information, resulting in a dense graph with valued edges. Remove noise to detect significant relationships corresponds to estimate an unobserved underlying binary network.

To this end, we develop a multiple testing procedure that incorporates the graph topology : the interaction structure of the underlying graph is learned with a Noisy version of the Stochastic Block Model. Parameter estimates and a node clustering is then provided via a variational expectation-maximization algorithm. It can be shown that our procedure asymptotically mimics the oracle procedure that controls the false discovery rate (FDR, average proportion of errors among the detected edges) while maximizing the true discovery rate (TDR, average proportion of recovered true edges). Numerical experiments illustrate the performance of our test procedure and show that it outperforms classical methods.

References

[1] Rebafka T., Roquain E., Villers F. Graph inference with clustering and false discovery rate control. *arXiv e-prints*, 2021.