Master Thesis Offer:
Screening for Graph Learning

Supervisor: Benjamin Girault
E-mail: benjamin.girault@inria.fr
Location: Inria MALICE project-team, Laboratoire Hubert Curien, Saint-Étienne, France
Level: Master 1 or 2 / 2nd or 3rd year of engineering school
Gratuity: ≃ 600 euros/month
Keywords: Graph Learning; Graph signal processing; Maximum likelihood estimator

Motivations and Description

Faced with multivariate data, a natural goal is to model relations between the variables of the system to leverage their inter-dependencies. For example, using correlations between close sensors in a sensor network is advantageous over ignoring them and having one model per sensor. Such a model is best described by a graph where nodes are variables and edges connect nodes having dependencies. Ultimately, the *graph signal processing* toolbox can be leveraged on such a graph to study the data.

However, such a graph depends on the task at hand, and, as such, is not given by the dataset. We recently introduced a novel method using maximum likelihood to derive edge weights and vertex weights [1]:

$$\min_{\mathbf{Q},\mathbf{L}} -\operatorname{logdet}(\mathbf{Q}+\mathbf{L}) + \operatorname{tr}\left[(\mathbf{Q}+\mathbf{L})\boldsymbol{\Sigma}\right]$$

where $\mathbf{L}$ is the graph Laplacian matrix given by edge weights, $\mathbf{Q}$ is the diagonal matrix of vertex weights and $\boldsymbol{\Sigma}$ is the data covariance matrix. This new method is a proof of concept that shows great promise, especially with respect to sparsity of the obtained graph compared to the state of the art [3]. The goal of this internship is threefold: i) implement variable screening [2] to speed up the computations, ii) run experiments on synthetic and real datasets, and compare our method to the state of the art, and iii) explore alternative cost functions.

Expected Results

- Implement variable screening [2] to speed up the computations
- Run experiments on synthetic and real datasets (in Python)
- Comparison metrics with the state of the art
- Explore alternative cost functions

References

[1] B. Girault, E. Pavez, and A. Ortega. Joint graph and vertex importance learning. In *2023 31st European Signal Processing Conference (EUSIPCO)*, pages 1858–1862, 2023.

[2] C. Herzet, C. Elvira, and H.-P. Dang. Region-free Safe Screening Tests for ℓ_1-penalized Convex Problems. In *2022 30th European Signal Processing Conference (EUSIPCO)*, page 2061–2065. IEEE, 2022.

[3] E. Pavez. Laplacian constrained precision matrix estimation: Existence and high dimensional consistency. In *Proc. of the 25th International Conference on Artificial Intelligence and Statistics*, Mar 2022.