

Lagrangian relaxation for tree-structured covariance estimation

Xiyang Dai, Zebao Gao and Hao Zhou
Dept. of Computer Science

1. TREE-STRUCTURED COVARIANCE ESTIMATION

Understanding similarities in expression profiles from a set of samples, e.g., multiple brain regions, is valuable to help understand many biological processes. A natural way to analyze this kind of similarities is to model a tree-structured covariance estimation. However, this kind of models turns out to be a non-convex optimization problem shown in recent paper from Hector et al [1]. In the paper, they have proposed a mixed-integer programming (MIP) approach to solve this non-convex optimization problem. Specifically, they estimate a covariance matrix from observations of p continuous random variables encoding a stochastic process over a tree with p leaves. To formulate the estimation problem as instances of well-studied numerical optimization problems, they used linear combinations of rank-one matrices indicating object partitions. Although they have shown great performance using MIP, this non-convex optimization problem for tree-structure covariance estimation is still far from solved.

2. LAGRANGIAN RELAXATION

Recently, there is an increasing trend in applying Lagrangian relaxation method to solve non-convex optimization problems. It has been successfully applied to several NLP inference problems such as Part-of-speech tagging [2]. But more studies are still desirable to conduct to evaluate its practicability in the tree-structured covariance estimation of our Bioinformatics problems.

3. OUR PROJECT

In this project, we will try to see whether Lagrangian relaxation method can be applied to solve this kind of non-convex optimization problem, and try to apply this tree-structured estimation method to Bio-related topics. The main challenges of our project include:

- 1 understanding the existing tree-structured covariance estimation problem;
- 2 understanding the Lagrangian relaxation method;
- 3 try to apply the method from NLP field to bio-related topics;
- 4 implementation of these methods.

REFERENCES

- [1] Héctor Corrada Bravo, Stephen J. Wright, Kevin H. Eng, Sunduz Keles, and Grace Wahba. Estimating tree-structured covariance matrices via mixed-integer programming. In *AISTATS*, JMLR Proceedings, pages 41–48, 2009.
- [2] Alexander M. Rush and Michael Collins. A tutorial on dual decomposition and lagrangian relaxation for inference in natural language processing. *Journal of Artificial Intelligence Research*, 2012.