

Learning the structure of a Markov Random Field

Lee Richert (ECE), Raghav Kuppan (ECE), Yuanda Zhu (ECE)

1 Abstract

Learning the structure of a Markov Random Field from independent and identically distributed samples is an important problem. Unfortunately, calculation of the partition function is intractable for most graphs. Trace Lasso regularization of a pseudo-likelihood has been proposed as a means of learning the structure of probabilistic graphical models from samples. The Trace Lasso exhibits less volatility in response to correlation between random variables than Lasso while also promoting sparsity. We plan to adapt Trace Lasso for the purpose of learning the structure of a Markov Random field, using a maximum pseudo-likelihood approach.

2 Introduction

2.1 Defining the Ising Model

An Ising Model is a special case of a pairwise Markov Random Field where each vertex in the graph takes on values in $\{-1,1\}$. The node potentials and edge potentials for an Ising model have very simple expressions thereby giving us a distribution of the form

$$\mathbb{P}_{\theta}^*(x) = \frac{1}{\mathbb{Z}(\theta^*)} \exp \sum_{(s,t) \in E} \theta_{st}^* x_s x_t$$

The Ising Model was proposed as a mathematical model for ferromagnetism in statistical mechanics but is also used in other applications such as Computer Vision and Neuroscience.

3 Prior Work on Model Selection

Learning the structure of Ising models is a challenging problem. For a general graph with p nodes of degree at most d , an exhaustive search across all possible edges takes around p^d computations. This is the time required to exhaustively search over all possible neighborhoods of a node and for each node test whether conditional independence assumptions hold. As d grows, the computational cost becomes untenable, so algorithms with lower

computational complexity are desired. In general, efficient algorithms for structural learning either restrict the graph structure or the nature of the interactions between the nodes. One possible assumption of the second kind is the Correlation Decay Property. A graphical model is said to have the correlation decay property if any two variables are asymptotically independent as the graph distance between them increases. This property holds for Ising models in certain real-world problems such as the ferromagnetic model in a high temperature regime. Alternative methods that do not explicitly require the Correlation Decay Property are usually based on Convex Optimization.

Ravikumar et. al [?] study the problem of signed edge recovery on Ising models and propose an l_1 -regularized logistic regression approach to recover the signed edges in the graph. They establish sufficient conditions on the sample size n , dimension p , and maximum neighborhood size d , to get a consistent estimator. With this assumption, the structure of any bounded degree graph can be recovered with high probability once $n/\log(p)$ is sufficiently large. This way, the signed neighborhood of every node can be estimated by optimizing a log pseudo-likelihood function with an ℓ_1 penalty. This way the entire set of signed edges of the graph can be recovered whereby the structure of the Ising model has been found. his technique is demonstrated on four-nearest neighbor lattices, eight-nearest neighbor lattices and a star graph as well as on a class of graphs with unbounded maximum neighborhood size, with the results being consistent with the theoretical conjectures

3.1	Learning Ising Models from Complete Data
3.2	Pseudolikelihood Functions
3.3	Regularization
3.3.1	Lasso
3.3.2	Group Lasso
3.3.3	Elastic Net
3.3.4	Trace Lasso
4	Approach
4.1	Parameter Selection
4.2	Applying Methods to Categorical Data
4.3	Learning Graphical Models for Classification
5	Datasets
6	Evaluating Performance
7	Logistics
8	Experiments
8.1	Synthetic Data
8.2	Congressional Voting Data
8.3	Mushroom Data
9	Conclusions
10	Acknowledgements