

Network structure from rich but noisy data

Natalia Garcia Martin & Maud Lemerrier

2018-10-14

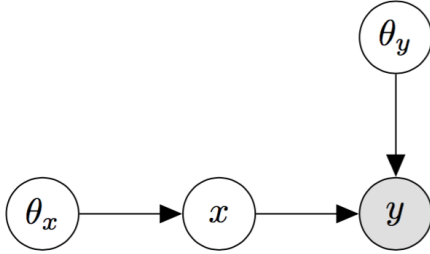
Abstract

We have implemented an R package to estimate the parameters of exponential random graph models (ERG models), in the case where the data consists of noisy observations of the underlying hidden structure. The first part of this report focusses on the Bernoulli (or Erdos Renyi) model, performing inference via the Expectation Maximization algorithm, while the second part explores other ERG models, for which Bayesian parameter estimation is more challenging.

1 Introduction

1.1 Motivation

The true network structure is drawn from $P(A|\theta)$. We will later constrain this distribution to belong to the exponential random graph models family. The observations of the network are supposed to be noisy. The network structure and the observations are related to one another by $P(data|A, \theta)$. Our aim is to infer the parameters θ through the posterior distribution $P(\theta|data)$. This is challenging, since it involves marginalising over the hidden variables, leading to intractable computations.



$$\begin{aligned} P(\theta|data) &\propto P(data|\theta)P(\theta) \\ &\propto \sum_A P(data, A|\theta)P(\theta) \\ &\propto \sum_A P(data|A, \theta)P(A|\theta)P(\theta) \end{aligned}$$

1.2 Models

1.2.1 Bernoulli

Let ρ be the probability of an edge in any position. Then, the Bernoulli model is defined by:

$$P(A|\rho) = \prod_{i < j} \rho^{A_{i,j}} (1 - \rho)^{1 - A_{i,j}}$$

1.2.2 Generalisation

$$P_{\theta}(A = a) = \exp \left(\sum_{i=1}^k \theta_i T_i(a) - c(\theta) \right)$$

2 Inference for the Bernoulli model

2.1 Inference via the Expectation Maximization algorithm

2.2 Experiments and results

2.2.1 Tests on simulated data

2.2.2 Tests on real data

3 Extension to other Exponential Random Graph Models

TO DO: explain why we will not use the EM algorithm.

The aim of this section is to discuss how a bayesian estimation of the parameters $\theta = \{\theta_x, \theta_y\}$ can be performed. In our case, we cannot use a vanilla Metropolis Hastings algorithm with $p(\theta|y)$ as a target distribution, for the following reasons:

- The normalising constant $Z(\theta)$ is intractable. This problem is would also arise in the case where the network is completely observed.
- The likelihood appearing in the acceptance ratio is intractable, since we would have to sum over all the states that the hidden variable x can take, with x a binary vector of size n .

$$\alpha_{MH}(\theta, \theta') = 1 \wedge \frac{p(y|\theta')p(\theta')q(\theta|\theta')}{p(y|\theta)p(\theta)q(\theta'|\theta)}$$

$$p(y|\theta) = \sum_x p(y|\theta, x)p(x|\theta)$$

3.1 Experiments