Bayesian networks, Kullback-leibler and topology

Ait-taleb Nabil *

Abstract

In this paper, I will propose a topology allowing to measure a neighborhood for the Bayesian networks. This topology will correspond to a Kullback-Leibler distance ratio and will allow to know the distance between a current Bayesian network and a Bayesian network having a chain rule. This topology applied to Bayesian networks will be normalized and will therefore vary from 0 to 1. The value 0 will correspond to a Bayesian network with a chain rule and the value 1 to a Bayesian network without edges.

*Corresponding author: nabiltravail1982@gmail.com

1 Introduction

In this paper we will use the results obtained in the report [6] page 21 in order to propose a topology corresponding to a Kullback-liebler distance ratio.

This topology will vary from 0 to 1 and will give 0 for a Bayesian network having a chain rule and 1 for a Bayesian network without edges.

From a fixed neighborhood ϵ around the chain rule, we will propose an algorithm allowing to select an optimal Bayesian network.

2 Bayesian networks, Kullback-leibler and topology

In what follows we will use a Kullback-leibler distance ratio as a topology measuring the neighborhood of a current Bayesian network $\mathcal{B}$ to the chain rule of the directed acyclic graph:

$$\boxed{0 \le \frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B})]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]} \le 1}$$

Where $\mathcal{B}^C$ will correspond to the chain rule of the Bayesian network and $\mathcal{B}^R$ to the Bayesian network without edges.

The inequality is based on the paper [6] page 21.

The lower bound 0 corresponds to the Bayesian network which is a chain rule and the upper bound 1 corresponds to the Bayesian network whithout edges: $\frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B}^C)]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]} = 0$ and $\frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B}^R)]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]} = 1$

1. The goal of the algorithm is to start from a Bayesian network without edges $\mathcal{B}^R$.
2. Then we add the edges producing the strongest variations of conditional entropy in order to produce the strongest oriented dependencies without producing cycles in the graph.
3. The current Bayesian network $\mathcal{B}$ then has a ratio $\frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B})]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]}$ which will decrease by adding the edges.
4. When the current ratio $\frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B})]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]}$ is in a good neighborhood of the chain rule $\mathcal{B}^C$:

$$0 \le \frac{D_{KL}[P_X(\vec{x}|\mathcal{B}^C)||P_X(\vec{x}|\mathcal{B})]}{D_{KL}[P_X(\vec{x}|\mathcal{B}^C||P_X(\vec{x}|\mathcal{B}^R)]} \le \epsilon$$

   we select the most likely Bayesian network $\mathcal{B}$.

   A good neighborhood is fixed for ϵ equal to 1%: $\epsilon = 0.01$

3 Conclusion

In this paper, we proposed a topology based on the Kullbak-leibler distance and allowing the selection of an optimal Bayesian network. The topology allows to put two bounds in which a current Bayesian network moves.

Starting from a Bayesian network without edges, the goal of the algorithm was then, by adding the edges, to find the Bayesian network closest to the lower bound corresponding to a chain rule with a neighborhood value ϵ set to 0.01.

[1]Elements of information theory. Author: Thomas M.Cover and Joy A.Thomas. Copyright 1991 John Wiley and sons.

[2]Optimal stastical decisions. Author: Morris H.DeGroot. Copyright 1970-2004 John Wiley and sons.

[3]Matrix Analysis. Author: Roger A.Horn and Charles R.Johnson. Copyright 2012, Cambridge university press.

[4]Causality: Models, reasoning and inference. Author:Judea Pearl .Copyright 2000, Cambridge university press.

[5]Bayesian Network and information Theory. Year:2022, Published:viXra, Category: Artificial Intelligence. Author: Ait-Taleb Nabil.

[6] Information theory applied to Bayesian network for learning continuous Data Matrix. Year:2021, Published:viXra, Category: Artificial Intelligence. Author: Ait-Taleb Nabil.