Quotient Normalized Maximum Likelihood Criterion for Learning Bayesian Network Structures (/pdf?id=S1tHmlnhe)

Blinded names

31 Mar 2017 (modified: 01 Apr 2017) UAI 2017 readers: UAI 2017, UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper123 Authors Original (/forum?id=BytrQen2l)

Student paper: No

Abstract: We introduce an information theoretic criterion for Bayesian network structure learning which we call quotient normalized maximum likelihood (qNML). In contrast to the closely related factorized normalized maximum likelihood criterion, qNML satisfies the property of score equivalence. It is also decomposable and completely free of adjustable hyperparameters. For practical computations, we identify a remarkably accurate approximation proposed earlier by Szpankowski and Weinberger. Experiments on both simulated and real data demonstrate that the new criterion leads to parsimonious models with good predictive accuracy.

TL;DR: A new model selection criterion for Bayesian network structure learning

Paperhash: names|quotient_normalized_maximum_likelihood_criterion_for_learning_bayesian_network_structures

Authorids: auai.org/UAI/2017/Paper123/Authors

Keywords: Bayesian networks, structure learning, model selection

Subject areas: Learning: Structure Learning, Models: Bayesian Networks

4 Replies

Add **Open Comment**

This paper may be accepted unless more excellent papers occupy the slots.

UAI 2017 Paper123 AnonReviewer3

15 May 2017 UAI 2017 Paper123 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper123 Authors

Rating: 7: Good paper, accept

Review: This paper proposes a quotient version of NML based BNSL.
The fNML scores are replaced by the quotient NML ones.
I understand the merits of the quotient strategy that the author claims:
the score is equivalent (Sect. 3.1) and the consistent (Sect. 3.2) and other merits of NML are still available the three theorems are relatively easy to derive.

Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

Add **Open Comment**

new scoring metric not performing well

UAI 2017 Paper123 AnonReviewer1

15 May 2017 UAI 2017 Paper123 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper123 Authors

Rating: 4: Ok but not good enough - rejection

Review: The authors propose a new approximation of the normalized maximum likelihood (NML). The existing factorized NML sometimes yields overly complex models, plus it is not score equivalent. The new qNML, although simple, is both score equivalent and consistent, which are nice properties to have. The quotient idea is from [Suzuki 16] if I'm not mistaken.

In the empirical results, the qNML score is compared to BDeu, BIC, and fNML. The authors seem to have a different perspective of the empirical results. For me, I am not impressed by the runner-up tendencies of qNML. BIC is claimed to be biased toward simplicity. But BIC models seem to have best predictive log losses for almost half of the models, and BIC is seldom worst among all scoring metrics. Plus, BIC models are often the simplest. It is probably a matter of personal taste. But I would prefer a method that is at least able to beat other methods now and then.

Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

Add **Open Comment**

The paper is good from the aspect of theory but does not show strong evidence for its usefulness in real application

UAI 2017 Paper123 AnonReviewer4

15 May 2017 UAI 2017 Paper123 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper123 Authors

Rating: 6: Marginally above acceptance threshold

Review: The authors of this paper propose a new scoring criterion called quotient normalized maximum likelihood (qNML) for structure learning in Bayesian networks (BNs). The authors theoretically prove some good properties of qNML including score equivalency (for Markov equivalent networks) and consistency. Finally the authors empirically show the performance of qNML versus other commonly used scoring criteria including BDeu and BIC.

Overall, I think that the paper demonstrates both novelty and importance from the aspect of theory but does not show strong evidence for its advantages over other commonly used scoring criteria in real application.

On one hand, the authors use the trick mentioned after Eq. (8) to develop qNML, which advances fNML (factorized normalized maximum likelihood) scoring criteria and is shown to have the additional score-equivalent property. The computation costs for qNML also equal the cost of other commonly used scoring criteria. I really appreciate these developments.

On the other hand, I think that the authors do not provide strong evidence for the usefulness of the developed qNML in Section 4, which largely affects the strength of the paper. In the following some details are list for the further improvement of the paper:

(1). The authors do not specify the equivalent sample size used for BDeu during the comparison. Is the equivalent sample size always equal to 1 in all the comparison settings (including Figure 1)? Could some simple rule of thumb be found to adjust the equivalent sample size according to the characteristics of the data (such as the size of the data) so that BDeu can overcome its drawback shown in Figure 1?
Even without any adjustment of the equivalent sample size, Figures 2 and 3 show that BDeu significantly outperforms the proposed qNML when the parameter generation mechanism matches the assumptions
of the BDeu-score, and performs very similarly with qNML when the parameters are drawn from the Jeffreys' prior. As a result, I cannot clearly see the advantages of the proposed qNML over BDeu. I would also like to see the performance comparison under other parameter generation mechanisms. In addition, could the performance for the specified setting (i.e. 5 node Bayesian network structures with 4/7 edges) generalize well (such as to larger networks)?

(2). Ranks remove the useful information about the actual score difference. Since both qNML and BDeu return the posterior probability, the score of an underlying BN can be reported across different sample sizes for both criteria. When the sample size is large, a better scoring criterion should give a larger value (probability) to the underlying BN. By this way, BNs with large number of variables can be examined for comparison.

(3). I are concerned about whether the error from the approximation from Eq. (7) will accumulate and eventually affect the performance of qNML for large BNs. Could the authors provide some evidence to remove my concern?

Confidence: 3: The reviewer is fairly confident that the evaluation is correct

Add **Open Comment**

The authors propose a new scoring function for learning BN structures. The paper is nice.

UAI 2017 Paper123 AnonReviewer2

10 May 2017 UAI 2017 Paper123 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper123 Authors

Rating: 7: Good paper, accept

Review: The authors propose a variant of NML called qNML. They prove that it is both consistent and score equivalent. Finally, they show in the experimental section that this new score is a good candidate for finding generating structures as well as BNs well suited for prediction tasks.

The paper is well written and proposes a novel score which seems to have good properties. The mathematical part is sound. But, maybe, it would be useful to say explicitly in Theorem 3 that the C connected components are also tournaments. This is mentioned before the theorem but not within its statement. In my opinion, this theorem and, more generally Subsection 3.3, is not very useful: actually, in practical applications, it seems very unlikely that the set of conditional independence statements satisfied by probability distributions lead to graphical structures defined by C mutually independent tournaments.

In the experimental section, Subsection 4.1 could be more convincing using classical benchmark Bayesian networks rather than randomly generated structures because, if care is not taken, the latter are usually very far from structures that can be encountered in practical situations and this can bias the results. In addition, as the proof of Theorem 2 is based on the fact that qNML is asymptotically equivalent to the BIC score, it may be wondered what is the minimal size of the database to get both scores more or less the same. In Subsection 4.2, it is mentioned that "BIC often wins with small sample sizes, but it performs worst for the large sample sizes". Maybe, here, some explanations would be of interest: intuitively, I would have thought that qNML would be of interest for small databases because, for large ones, asymptotically, it is equivalent to BIC. Yet, Subsection 4.2 seems to suggest the converse.

Overall, the paper presents an interesting contribution.

Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

Add **Open Comment**