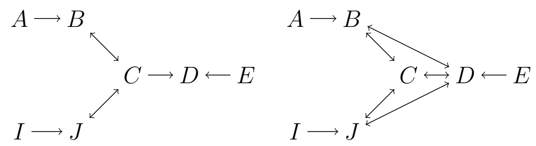
LEARNING MULTIVARIATE REGRESSION CHAIN GRAPHS UNDER FAITHFULNESS: ADDENDUM

JOSE M. PEÑA
ADIT, IDA, LINKÖPING UNIVERSITY, SE-58183 LINKÖPING, SWEDEN
JOSE.M.PENA@LIU.SE

The correctness of our algorithm in the main text lies upon the assumption that p is faithful to some MVR CG. This is a strong requirement that we would like to weaken, e.g. by replacing it with the milder assumption that p satisfies the composition property. Correct algorithms for learning directed and acyclic graphs (a.k.a. Bayesian networks) under the composition property assumption exist (Chickering and Meek, 2002; Nielsen et al., 2003). We have recently developed a correct algorithm for learning LWF CGs under the composition property (Peña et al., 2014). The way in which these algorithms proceed (a.k.a. score+search based approach) is rather different from that of our algorithm (a.k.a. constraint based approach). In a nutshell, they can be seen as consisting of two phases: A first phase that starts from the empty graph H and adds single edges to it until p is Markovian with respect to H, and a second phase that removes single edges from H until p is Markovian with respect to H and p is not Markovian with respect to any CG F such that $I(H) \subseteq I(F)$. The success of the first phase is guaranteed by the composition property assumption, whereas the success of the second phase is guaranteed by the so-called Meek's conjecture (Meek, 1997). Specifically, given two directed and acyclic graphs F and H such that $I(H) \subseteq I(F)$, Meek's conjecture states that we can transform F into H by a sequence of operations such that, after each operation, F is a directed and acyclic graph and $I(H) \subseteq I(F)$. The operations consist in adding a single edge to F, or replacing F with a triplex equivalent directed and acyclic graph. Meek's conjecture was proven to be true in (Chickering, 2002, Theorem 4). The extension of Meek's conjecture to LWF CGs was proven to be true in (Peña et al., 2014, Theorem 1). The extension of Meek's conjecture to AMP CGs was proven to be false in (Peña, 2012, Example 1). The example below shows that the extension of Meek's conjecture to MVR CGs does not hold either.

Consider the MVR CGs F and H below.



F H

We can describe $I(F)$ and $I(H)$ by listing all the separators between any pair of distinct nodes. We indicate whether the separators correspond to F or H with a superscript. Specifically,

- $\mathcal{S}^F_{AB} = \mathcal{S}^F_{BC} = \mathcal{S}^F_{CD} = \mathcal{S}^F_{DE} = \mathcal{S}^F_{IJ} = \mathcal{S}^F_{JC} = \varnothing$,
- $\mathcal{S}^F_{AC}$ = all the node sets that do not contain $\{B\}$,
- $\mathcal{S}^F_{AD}$ = all the node sets that do not contain $\{B\}$ or contain $\{C\}$,
- $\mathcal{S}^F_{AE}$ = all the node sets that do not contain $\{B, D\}$ or contain $\{C\}$,
- $\mathcal{S}^F_{AI}$ = all the node sets that do not contain $\{B, C, J\}$,
- $\mathcal{S}^F_{AJ}$ = all the node sets that do not contain $\{B, C\}$,

Date: 12:37, 05/02/14.

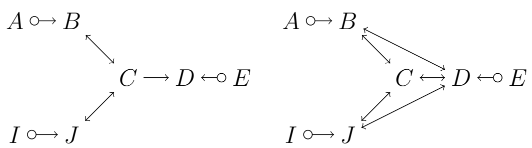
- $\mathcal{S}^F_{BD}$ = all the node sets that contain $\{C\}$,
- $\mathcal{S}^F_{BE}$ = all the node sets that do not contain $\{D\}$ or contain $\{C\}$,
- $\mathcal{S}^F_{BI}$ = all the node sets that do not contain $\{C, J\}$,
- $\mathcal{S}^F_{BJ}$ = all the node sets that do not contain $\{C\}$,
- $\mathcal{S}^F_{CE}$ = all the node sets that do not contain $\{D\}$,
- $\mathcal{S}^F_{CI}$ = all the node sets that do not contain $\{J\}$,
- $\mathcal{S}^F_{DI}$ = all the node sets that do not contain $\{J\}$ or contain $\{C\}$,
- $\mathcal{S}^F_{DJ}$ = all the node sets that contain $\{C\}$,
- $\mathcal{S}^F_{EI}$ = all the node sets that do not contain $\{D, J\}$ or contain $\{C\}$,
- $\mathcal{S}^F_{EJ}$ = all the node sets that do not contain $\{D\}$ or contain $\{C\}$.

Likewise,

- $\mathcal{S}^H_{AB} = \mathcal{S}^H_{BC} = \mathcal{S}^H_{BD} = \mathcal{S}^H_{CD} = \mathcal{S}^H_{DE} = \mathcal{S}^H_{DJ} = \mathcal{S}^H_{IJ} = \mathcal{S}^H_{JC} = \varnothing$,
- $\mathcal{S}^H_{AC}$ = all the node sets that do not contain $\{B\}$,
- $\mathcal{S}^H_{AD}$ = all the node sets that do not contain $\{B\}$,
- $\mathcal{S}^H_{AE}$ = all the node sets that do not contain $\{B, D\}$,
- $\mathcal{S}^H_{AI}$ = all the node sets that contain neither $\{B, C, J\}$ nor $\{B, D, J\}$,
- $\mathcal{S}^H_{AJ}$ = all the node sets that contain neither $\{B, C\}$ nor $\{B, D\}$,
- $\mathcal{S}^H_{BE}$ = all the node sets that do not contain $\{D\}$,
- $\mathcal{S}^H_{BI}$ = all the node sets that contain neither $\{C, J\}$ nor $\{D, J\}$,
- $\mathcal{S}^H_{BJ}$ = all the node sets that contain neither $\{C\}$ nor $\{D\}$,
- $\mathcal{S}^H_{CE}$ = all the node sets that do not contain $\{D\}$,
- $\mathcal{S}^H_{CI}$ = all the node sets that do not contain $\{J\}$,
- $\mathcal{S}^H_{DI}$ = all the node sets that do not contain $\{J\}$,
- $\mathcal{S}^H_{EI}$ = all the node sets that do not contain $\{D, J\}$,
- $\mathcal{S}^H_{EJ}$ = all the node sets that do not contain $\{D\}$.

Then, $I(H) \subseteq I(F)$ because $\mathcal{S}^H_{XY} \subseteq \mathcal{S}^F_{XY}$ for all $X, Y \in \{A, B, C, D, E, I, J\}$ with $X \neq Y$.

All the CGs that are triplex equivalent to F and H can be represented by the graphs $\mathcal{F}$ and $\mathcal{H}$ below, where a circle at the end of an edge represents an unspecified end, i.e. an arrowhead or nothing.



$\mathcal{F}$ $\mathcal{H}$

However, we cannot transform any CG in $\mathcal{F}$ into a CG in $\mathcal{H}$ as required by Meek's conjecture. To see it, note that the only modifications that we can perform to any CG in $\mathcal{F}$ is adding the edge $B \to D$ or the edge $J \to D$. This implies that $A \not\perp D$ or $J \not\perp D$ in the resulting CG, whereas $A \perp D$ and $J \perp D$ in any CG in $\mathcal{H}$.

References

Chickering, D. M. Optimal Structure Identification with Greedy Search. *Journal of Machine Learning Research*, 3:507-554, 2002.

Chickering, D. M. and Meek, C. Finding Optimal Bayesian Networks. In *Proceedings of 18th Conference on Uncertainty in Artificial Intelligence*, 94-102, 2002.

Meek, C. *Graphical Models: Selecting Causal and Statistical Models*. PhD thesis, Carnegie Mellon University, 1997.

Nielsen, J. D., Kočka, T. and Peña, J. M. On Local Optima in Learning Bayesian Networks. In *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence*, 435-442, 2003.

Peña, J. M. Learning AMP Chain Graphs under Faithfulness. In *Proceedings of the 6th European Workshop on Probabilistic Graphical Models*, 251-258, 2012.

Peña, J. M., Sonntag, D. and Nielsen, J. D. An Inclusion Optimal Algorithm for Chain Graph Structure Learning. In *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics*, 2014, accepted.