Inference of Cause and Effect with Unsupervised Inverse Regression

Abstract

We address the problem of causal discovery in the two-variable case, given a sample from their joint distribution. Since $X \rightarrow Y$ and $Y \rightarrow X$ are Markov equivalent, conditional-independence-based methods [Spirtes et al., 2000, Pearl, 2009] can not recover the causal graph. Alternative methods, introduce asymmetries between cause and effect by restricting the function class (e.g., [Hoyer et al., 2009]).

The proposed causal discovery method, CURE, is based on the principle of independence of causal mechanisms [Janzing and Schölkopf, 2010]. For the case of only two variables, it states that the marginal distribution of the cause, say $P(X)$, and the conditional of the effect given the cause $P(Y|X)$ are "independent", in the sense that they do not contain information about each other (informally $P(X)$ "$\perp\!\!\!\perp$" $P(Y|X)$). This independence can be violated in the backward direction: the distribution of the effect $P(Y)$ and the conditional $P(X|Y)$ may contain information about each other because each of them inherits properties from both $P(X)$ and $P(Y|X)$, hence introducing an asymmetry between cause and effect. For deterministic causal relations ($Y = f(X)$), all the information about the conditional $P(Y|X)$ is contained in the function f, so independence boils down to $P(X)$ "$\perp\!\!\!\perp$" f. Previous work formalizes the independence principle by specifying what is meant by independence. For deterministic non-linear relations, Janzing et al. [2012] and Daniusis et al. [2010] define independence as uncorrelatedness between $\log f'$ and the density of $P(X)$, both viewed as random variables. For non-deterministic relations, it is not obvious how to explicitly formalize independence between $P(X)$ and $P(Y|X)$. Instead, we propose an implicit notion of independence, namely that $p_{Y|X}$ cannot be estimated based on p_X (lower case denotes density). However, it may be possible to estimate $p_{X|Y}$ based on the density of the effect, p_Y.

In practice, we are given empirical data $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{y} \in \mathbb{R}^N$ from $P(X, Y)$ and estimate $p_{X|Y}$ based on $\mathbf{y}$ (intentionally hiding $\mathbf{x}$). The relationship between the observed $\mathbf{y}$ and the *latent* $\mathbf{x_u} \in \mathbb{R}^N$ is modeled by a Gaussian Process (GP): $p(\mathbf{y}|\mathbf{x_u}, \boldsymbol{\theta}) = \mathcal{N}(\mathbf{y}; \mathbf{0}, K_{\mathbf{x_u},\mathbf{x_u}} + \sigma_n^2 I_N)$ (this can be alternatively seen as a single output GP-LVM). Then, the required conditional $p_{X|Y}$ is estimated as $\hat{p}^{\mathbf{y}}_{X_u|Y} : (x_u, y) \mapsto p(x_u|y, \mathbf{y})$, with $p(x_u|y, \mathbf{y})$ estimated by marginalizing out the latent $\mathbf{x_u}$ and $\boldsymbol{\theta}$ (GP hyperparameters).

CURE infers the causal direction by using the procedure above two times: one to estimate $p_{X|Y}$ based only on $\mathbf{y}$ and another to estimate $p_{Y|X}$ based only on $\mathbf{x}$. If the first estimation is better, $X \rightarrow Y$ is inferred. Otherwise, $Y \rightarrow X$. CURE was evaluated on synthetic and real data and often outperformed existing methods. On the downside, its computational cost is comparably high. This work was recently published at AISTATS 2015 [Sgouritsa et al., 2015].

References

P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search.* MIT Press, 2nd edition, 2000.

J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2nd edition, 2009.

P. Hoyer, D. Janzing, J. M. Mooij, J. Peters, and B. Schölkopf. Nonlinear causal discovery with additive noise models. In *Advances in Neural Information Processing Systems 21 (NIPS)*, 2009.

D. Janzing and B. Schölkopf. Causal inference using the algorithmic Markov condition. *IEEE Transactions on Information Theory*, 56:5168–5194, 2010.

D. Janzing, J. M. Mooij, K. Zhang, J. Lemeire, J. Zscheischler, P. Daniusis, B. Steudel, and B. Schölkopf. Information-geometric approach to inferring causal directions. *Artificial Intelligence*, 182-183:1–31, 2012.

P. Daniusis, D. Janzing, J. Mooij, J. Zscheischler, B. Steudel, K. Zhang, and B. Schölkopf. Inferring deterministic causal relations. In *Proceedings of the 26th Annual Conference on Uncertainty in Artificial Intelligence (UAI)*, 2010.

E. Sgouritsa, D. Janzing, P. Hennig, and B. Schölkopf. Inference of cause and effect with unsupervised inverse regression. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.