

طرح اولیه‌ی پروژه‌ی درس استنتاج علی (ویرایش دوم)

A Review on Additive Noise Models: Theory and Learning Algorithms

بهراد منیری
۹۵۱۰۹۵۶۴
bemoniri@live.com

دانشکده‌ی مهندسی برق - دانشگاه صنعتی شریف

۱ شرح مختصر طرح

در این پروژه قصد دارم تا مقالات مربوط به Additive Noise Models را به طور دقیق بررسی کنم. هدف من بررسی قضایا و الگوریتم‌های این حوزه، به خصوص الگوریتم‌های جدیدی مانند RESIT و مقایسه‌ی عملکرد این الگوریتم‌ها در یادگیری ساختار علی داده‌های واقعی است. البته توجه من تنها به الگوریتم‌ها معطوف نخواهد بود. یکی از اهداف مهم من در این پروژه، مطالعه و بررسی دقیق اثبات‌های مربوط به «قابل شناسایی بودن علت، در مدل‌های ANM» است. این موارد در حالتی که Hidden Variable وجود داشته، علاوه بر حالتی که تمام متغیرها مشاهده می‌شوند مورد مطالعه قرار خواهند گرفت. در سال‌های اخیر کار زیادی بر روی SCM ها با حلقه‌ی فیدبک و متغیر پنهان انجام می‌شود، مقاله‌ی [۶] مقاله‌ای بسیار جدید در این حوزه است.

۲ مقالات

مقالات زیر، مقالاتی هستند که قصد دارم در طی این پروژه به بررسی آنها بپردازم.
تمرکز من، بیشتر بر سه مقاله‌ی آخر خواهد بود.

1. Hoyer, P., Janzing, D., Mooij, J., Peters, J. and Schölkopf, B. (2008). *Nonlinear causal Discovery with Additive Noise Models*, **Advances in Neural Information Processing Systems 21 (NIPS 2008)**.

در این مقاله، برای دو متغیر که SCM ای به شکل زیر دارند اثبات می‌شود که با دیتای *observational* می‌توان جهت درست علیّت را تشخیص داد. این مقاله در واقع اثبات می‌کند که چگالی احتمالات معدودی هستند که اجازه‌ی وجود ANM در دو جهت را می‌دهند. در نهایت نیز روشی برای تشخیص این جهت درست ارائه شده و این روش بر روی داده‌های واقعی آزمایش می‌شود.

$$\begin{cases} X = N_x \\ Y = f(X) + N_y \end{cases} \quad N_x \perp\!\!\!\perp N_y \quad (1)$$

2. Zhang, K. and Hyvärinen, A. (2009). *On the Identifiability of the Post-Nonlinear Causal Model*, **UAI '09 Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence**, pp. 647-655.

در این مقاله، کار Hoyer et al. (2008) به SCM هایی به شکل

$$\begin{cases} X = N_x \\ Y = g(f(X) + N_y) \end{cases} \quad N_x \perp\!\!\!\perp N_y \quad (2)$$

تعمیم می‌یابد و روشی برای یافتن توابع f و g از دیتای مشاهداتی ارائه می‌شود. در نهایت نیز این روش بر روی دیتای واقعی آزمایش می‌شود.

3. Peters, J., Janzing, D. and Schölkopf B. (2010). *Identifying Cause and Effect on Discrete Data using Additive Noise Models*, **Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics**, PMLR 9, pp. 597-604.

این مقاله مسئله‌ای مشابه با مسئله‌ی (2008) Hoyer et al. را برای داده‌های گسسته بررسی می‌کند و نشان می‌دهد که در این حالت نیز توابع چگالی احتمال مشترک محدودی هستند که برای آنها ANM هایی در دو جهت وجود دارد. در ادامه الگوریتمی بهینه برای تشخیص علت با دیتای محدود ارائه شده و با داده‌های واقعی آزمایش می‌شود.

4. Peters, J. , Mooij, J., Janzing, D., and Schölkopf, B. (2014). *Causal Discovery with Continuous Additive Noise Models*, **Journal of Machine Learning Research** 15, pp. 2009-2053.

این مقاله در ابتدا مروری بسیار جامع بر مقالات قبل انجام می‌دهد و الگوریتمی عملی (و بهینه) به نام RESIT ارائه می‌دهد. این روش مبتنی بر Independence Score است. در این الگوریتم، هزینه‌ای برای افزایش تعداد یال‌ها قرار داده شده است تا در نهایت گرافی Causal Minimal به دست بیاید.

این مقاله دو الگوریتم RESIT را با روش GES و الگوریتم PC بر روی داده‌های واقعی آزمایش کرده و ادعا می‌کند که RESIT از دیگر الگوریتم‌ها موفق‌تر است.

5. Mooij, J., Janzing, D., Heskes, T., Schölkopf, B. (2011) *On Causal Discovery with Cyclic Additive Noise Models*, **Advances in Neural Information Processing Systems** 24 (NIPS 2011).

در این مقاله اثبات می‌شود که در حالت دو متغیره، با وجود حلقه‌ی فیدبک و با فرض نویز گاوسی و با فرض ANM، جهت درست علی قابل شناسایی است و الگوریتمی برای یادگیری این جهت از دیتای مشاهداتی ارائه می‌کند. در این روش باید رگرسیون غیرخطی با تابع هزینه‌ای مشابه Nonlinear ICA انجام دهیم. این مقاله، در آخر نیز این روش را بر روی داده‌های واقعی آزمایش می‌کند.

6. Forré, J., Mooij, J. (2018) *Constraint-based Causal Discovery for Non-Linear Structural Causal Models with Cycles and Latent Confounders*, **UAI '18 Proceedings of the 34th Conference on Uncertainty in Artificial Intelligence**.

این مقاله، مقاله‌ی بسیار جدیدی است که مدلی جامع برای علّیت ارائه می‌کند که در آن متغیرهای پنهان، حلقه‌های فیدبک و توابع غیرخطی وجود دارند. مفهوم σ - separation را به عنوان تعمیمی از d - separation تعریف می‌کند و از آن برای بررسی حلقه‌های فیدبک استفاده می‌کند. در ادامه الگوریتمی برای یادگیری ساختارهایی با متغیرهای پنهان و همچنین حلقه‌های فیدبک پیشنهاد می‌کند و آن را بر روی داده‌های واقعی آزمایش می‌کند.