

A Theory of Inferred Causation

Modelos Gráficos Causales

Sebastián Bejos

8 de julio de 2019

Instituto Nacional de Astrofísica Óptica y Electrónica

El Marco de Trabajo para el Descubrimiento Causal

- Se considera a la tarea del descubrimiento causal un **juego de inducción que los científicos juegan contra la Naturaleza**.
- La naturaleza posee mecanismos causales estables que, en un nivel detallado de descripciones, son **relaciones funcionales deterministas entre variables, algunas de las cuales son inobservables**.
- Estos mecanismos están organizados en forma de una **estructura acíclica**, que el científico intenta identificar a partir de las **observaciones disponibles**.

Definición (Estructura Causal)

Una estructura causal de un conjunto de variables \mathbf{V} es un grafo dirigido acíclico (DAG) en el que cada nodo corresponde a un elemento de \mathbf{V} y cada enlace representa una relación funcional directa entre las variables correspondientes.

Definición (Modelo Causal)

Un **modelo causal** es un par ordenado $M = (D, \Theta_D)$ que consiste de una **estructura causal** D y un **conjunto de parámetros** Θ_D compatibles con D . Los parámetros Θ_D asignan una función $x_i = f_i(pa_i, u_i)$ a cada $X_i \in \mathbf{V}$ y una medida de probabilidad $P(u_i)$ para cada u_i , donde \mathbf{PA}_i son los padres de X_i en D y donde cada U_i es una perturbación aleatoria que se distribuye de acuerdo con $P(u_i)$ independientemente de todas las demás u s.

Preferencia de Modelos (Occam's razor)

Siguiendo las normas estándar de inducción científica, es razonable descartar cualquier teoría para la cual encontremos una teoría más simple y menos elaborada que sea igualmente consistente con los datos.

Definición (Causalidad Inferida (Preliminar))

*Se dice que una variable X tiene una **influencia causal** en una variable Y , si existe una trayectoria dirigida de X a Y en **toda estructura minimal** consistente con los datos.*

Definición (Estructura Latente)

Una **estructura latente** es un par ordenado $L = (D, O)$, donde D es una estructura causal sobre \mathbf{V} y donde $O \subseteq \mathbf{V}$ es un conjunto de variables observables.

Definición (Preferencia de Estructura)

Una estructura latente $L = (D, O)$ es preferida a otra $L' = (D', O)$ (lo cual se denota como $L \preceq L'$) si y solo si D' puede imitar a D sobre O .

Esto es, si y solo si para todo Θ_D existe un $\Theta'_{D'}$, tal que

$$P_{[O]}((D', \Theta'_{D'})) = P_{[O]}((D, \Theta_D)).$$

Definición

Dos estructuras latentes son equivalentes, lo que se denota como $L' \equiv L$, si y solo si $L \preceq L'$ y $L \succeq L'$

Definición (Minimalidad)

Una estructura latente L es minimal con respecto a una clase \mathcal{L} de estructuras latentes si y solo si no existe un miembro de \mathcal{L} que sea estrictamente preferido a L . Esto es, si para todo $L' \in \mathcal{L}$ se tiene que $L \equiv L'$ siempre que $L' \preceq L$.

Definición (Consistencia)

Una estructura latente $L = (D, \mathbf{O})$ es consistente con una distribución \hat{P} sobre \mathbf{O} si D puede acomodar algún modelo que genere \hat{P} . Esto es, si existe una parametrización Θ_D tal que $P_{[\mathbf{O}]}(D, \Theta_D) = \hat{P}$.

Claramente, una condición necesaria (y algunas veces suficiente) para que L sea consistente es que L puede dar cuenta de todas las dependencias incorporadas en \hat{P} .

Definición (Causalidad Inferida)

Dada \hat{P} , una variable C tiene una influencia causal sobre otra variable E si y solo si existe una trayectoria dirigida de C a E en toda estructura latente minimal consistente con \hat{P}

Definición (Estabilidad)

Sea $I(P)$ el conjunto de todas las relaciones de independencias condicionales en P . Un modelo causal $M = (D, \Theta_D)$ genera una distribución estable si y solo si $P((D, \Theta_D))$ no contiene independencias extrañas. Esto es, si y solo si $I(P((D, \Theta_D))) \subseteq I(P((D, \Theta'_D)))$ para cualquier conjunto de parámetros Θ'_D