A Calculus of Intervention

Modelos Gráficos Causales

Sebastián Bejos, Ivan Feliciano

11 de junio de 2019

Instituto Nacional de Astrofísica Óptica y Electrónica

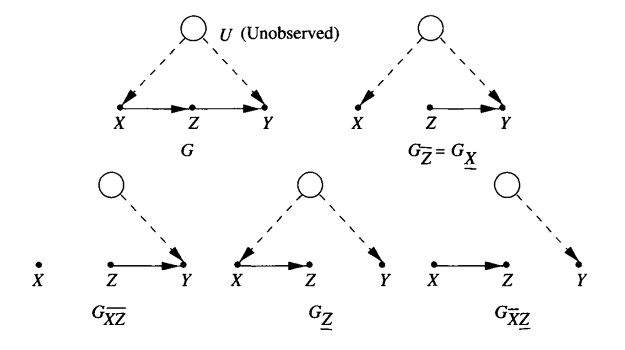
Introducción

- Esta sección establece un conjunto de reglas de inferencia que proporcionan un **método sintáctico para derivar (o verificar) afirmaciones sobre las intervenciones**.
- **Cada regla de inferencia respeta la interpretación del operador** $do(\cdot)$ como una intervención que modifica un conjunto seleccionado de funciones en el modelo subyacente.
- El conjunto de reglas de inferencia que emergen de esta interpretación es llamado **cálculo do**.

Introducción

- Se asume que **una estructura causal** G **es dada**, en donde algunos de los nodos son observables, mientras que otros permanecen sin ser observados.
- El objetivo es **facilitar la derivación sintáctica de las expresiones de efecto causal** de la forma $P(\mathbf{y} \mid do(\mathbf{x}))$, donde $\mathbf{X}$ y $\mathbf{Y}$ son subconjunto de variables observadas.
- Por "derivación" nos referimos a la reducción gradual de la expresión $P(\mathbf{y} \mid do(\mathbf{x}))$ a una **expresión equivalente que involucre probabilidades estándar de cantidades observadas**.
- Siempre que dicha reducción sea factible, se dice que el **efecto causal de X en Y es identificable**.

Notación preliminar



- Sea **X** y **Y** conjuntos disjuntos de nodos en la DAG causal G.
- Se denota por $G_{\overline{\mathbf{X}}}$ a la gráfica obtenida de borrar todos los arcos que apuntan a **X**, en G.
- Se denota por $G_{\underline{\mathbf{X}}}$ a la gráfica obtenida de borrar todos los arcos que surgen de **X**, en G.

Reglas del cálculo do

Teorema

Sea G una gráfica dirigida asociada a un modelo causal, y se $P(\cdot)$ la distribución de probabilidad inducida por ese modelo. Para cualesquiera conjuntos disjuntos de variables $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$ y $\boldsymbol{W}$, se tienen la siguientes reglas:

- ***Regla 1*** *(Inserción/Supresión de observaciones):*
  $P(\boldsymbol{y} \mid do(\boldsymbol{x}), \boldsymbol{z}, \boldsymbol{w}) = P(\boldsymbol{y} \mid do(\boldsymbol{x}), \boldsymbol{w}) \quad si\ (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{W})_{G_{\overline{\boldsymbol{X}}}}$.
- ***Regla 2*** *(Acción/intercambio de observación):*
  $P(\boldsymbol{y} \mid do(\boldsymbol{x}), do(\boldsymbol{z}), \boldsymbol{w}) = P(\boldsymbol{y} \mid do(\boldsymbol{x}), \boldsymbol{z}, \boldsymbol{w}) \quad si\ (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{W})_{G_{\overline{\boldsymbol{X}}\underline{\boldsymbol{Z}}}}$.
- ***Regla 3*** *(Inserción/Supresión de acciones):*
  $P(\boldsymbol{y} \mid do(\boldsymbol{x}), do(\boldsymbol{z}), \boldsymbol{w}) = P(\boldsymbol{y} \mid do(\boldsymbol{x}), \boldsymbol{w}) \quad si\ (\boldsymbol{Y} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{X}, \boldsymbol{W})_{G_{\overline{\boldsymbol{X}}, \overline{\boldsymbol{Z}(\boldsymbol{W})}}}$,
  donde $Z(W)$ es el conjunto de nodos de Z que no son ancestros de de ningún nodo en W, en $G_{\overline{\boldsymbol{X}}}$.

Reglas del cálculo do

- **Regla 1** (Inserción/Supresión de observaciones):
  $P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \quad \text{si } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{G_{\overline{\mathbf{X}}}}$.
- Esta regla, reafirma la separación-*d* como una prueba válida para la independencia condicional en la distribución resultante de la intervención $do(\mathbf{X} = \mathbf{x})$, de ahí la gráfica $G_{\overline{\mathbf{X}}}$.
- Esta regla se deduce del hecho de que al eliminar las ecuaciones del sistema no se introduce ninguna dependencia entre los restantes términos.

Reglas del cálculo do

- **Regla 2** (Acción/intercambio de observación):
  $P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{z}, \mathbf{w}) \quad$ si $(\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{G_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}}$.
- Esta regla, proporciona una condición para que una intervención externa $do(\mathbf{Z} = \mathbf{z})$ tenga el mismo efecto en **Y** que la observación pasiva $\mathbf{Z} = \mathbf{z}$.
- La condición equivale a $\{\mathbf{X} \cup \mathbf{W}\}$ bloquear todos los caminos de puerta trasera de **Z** a **Y** (en $G_{\overline{\mathbf{X}}}$), ya que $G_{\overline{\mathbf{X}}\underline{\mathbf{Z}}}$ retiene todas (y solo) dichas trayectorias.

Reglas del cálculo do

- **Regla 3** (Inserción/Supresión de acciones):
  $P(\mathbf{y} \mid do(\mathbf{x}), do(\mathbf{z}), \mathbf{w}) = P(\mathbf{y} \mid do(\mathbf{x}), \mathbf{w}) \quad \text{si } (\mathbf{Y} \perp\!\!\!\perp \mathbf{Z} \mid \mathbf{X}, \mathbf{W})_{G_{\overline{\mathbf{X}}, \overline{\mathbf{Z}(\mathbf{W})}}}$,
  donde $Z(W)$ es el conjunto de nodos de Z que no son ancestros de de ningún nodo en W, en $G_{\overline{\mathbf{X}}}$.
- Esta regla, proporciona condiciones para introducir (o eliminar) una intervención externa $do(\mathbf{Z} = \mathbf{z})$ sin afectar la probabilidad de $(\mathbf{Y} = \mathbf{y})$.
- La validez de esta regla se deriva, una vez más, de simular la intervención $do(\mathbf{Z} = \mathbf{z})$ mediante la eliminación de todas las ecuaciones correspondientes a las variables en $\mathbf{Z}$ (de ahí la gráfica $G_{\overline{\mathbf{X}\mathbf{Z}}}$).
- La razón para limitar la eliminación a los no antecesores de los nodes en $\mathbf{W}$ se proporciona en las pruebas de las Reglas 1-3 en Pearl(1995a).

Reglas del cálculo do

Corolario

Un efecto causal $q = P(y_1, \dots y_k \mid do(x_1), \dots do(x_m))$ *es identificable en un modelo caracterizado por una gráfica G si existe una secuencia finita de trasformaciones, cada una conforme a una de las tres reglas tres de inferencia (en el Teorema anterior), que reduce q en una expresión de probabilidad estándar (es decir, libre de ”sombreros”) que involucra cantidades observadas.*

Reglas del cálculo do

- Se ha demostrado que las Reglas 1-3 son completas, es decir, suficientes para obtener todos los efectos causales identificables (Shpitser y Pearl 2006a; Huang y Valtorta 2006)
- Las derivaciones simbólicas que usan la notación de sombrero son más convenientes que las derivaciones algebraicas que tienen como objetivo eliminar las variables latentes de las expresiones de probabilidad estándar.
- Sin embargo, la tarea de decidir si existe una secuencia de reglas para reducir una expresión de efecto causal arbitrario no se ha sistematizado y, por lo tanto, los criterios gráficos directos para la identificación son más deseables. Estos serán desarrollados en el Capítulo 4.

Un ejemplo de derivación simbólica de efectos causales

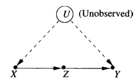


Figura 1

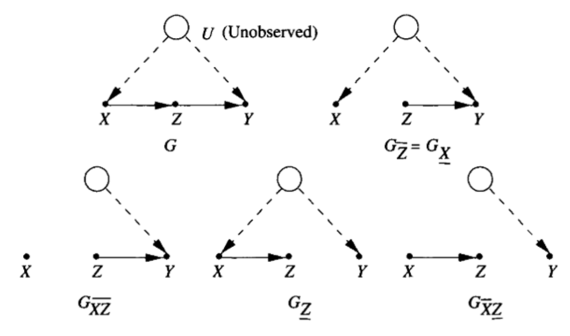


Figura 2: Subgrafos de *G* usados en la derivación de efectos causales.

Tarea 1: Calcular $P(z|do(x))$

Esta tarea se completa en un paso, ya que G cumple la condición de aplicabilidad para la Regla 2. Esto es, $X \perp\!\!\!\perp Z$ en $G_{\underline{X}}$ (porque el camino $X \leftarrow U \rightarrow Y \leftarrow Z$ está bloqueado por las aristas convergentes en Y), y podemos escribir

$$P(z|do(x)) = P(z|x)$$

Tarea 2: Calcular $P(y|do(z))$ i

Aquí no se puede aplicar la regla 2 (cambiar $do(z)$ por z) porque $G_{\underline{Z}}$ contiene un camino de puerta trasera desde Z a Y: $Z \leftarrow X \leftarrow U \rightarrow Y$. Nos gustaría bloquear este camino a través de la medición de variables (como X) que están en el camino. Esto involucra condicionar y sumar sobre todos los valores de X:

$$P(y|do(z)) = \sum_x P(y|x, do(z))P(x|do(z)).$$

$P(x|do(z))$ se puede calcular aplicando la regla 3;

$$P(x|do(z)) = P(x) \text{ si } (Z \perp\!\!\!\perp X)_{G_{\overline{Z}}},$$

porque X y Z son d-separados en $G_{\overline{Z}}$. (Intuitivamente, manipular Z no debe tener efecto en X, porque es descendiente de X en G).

Tarea 2: Calcular $P(y|do(z))$ ii

Para reducir $P(y|x, do(z))$, consultamos la Regla 2:

$$P(y|x, do(z)) = P(y|x, z) \text{ si } (Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$$

notando que X d-separa Z de Y en $G_{\underline{Z}}$.

Esto nos permite reescribir la ecuación como

$$P(y|do(z)) = \sum_x P(y|x, z)P(x) = E_x P(y|x, z),$$

el cual es un caso especial de la fórmula de la puerta trasera. La condición legitimadora $(Z \perp\!\!\!\perp Y|X)_{G_{\underline{Z}}}$, ofrece otra prueba gráfica para que un conjunto X sea suficiente para controlar factores de confusión (entre Y y Z).

Tarea 3: Calcular $P(y|do(x))$ i

Escribiendo

$$P(y|do(x)) = \sum_z P(y|z, do(x))P(z|do(x)).$$

Ya hemos reducido el término $P(z|do(x))$ a $P(z|x)$. Ninguna regla se puede aplicar para eliminar el operador $do()$ del término $P(y|z, do(x))$. Sin embargo, podemos añadir el operador a través de la Regla 2:

$$P(y|z, do(x)) = P(y|do(z), do(x))$$

ya que es aplicable la condición $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\underline{Z}}}$.

Tarea 3: Calcular $P(y|do(x))$ ii

Podemos ahora eliminar la acción $do(x)$ usando la regla 3, porque $(Y \perp\!\!\!\perp X|Z)_{G_{\overline{XZ}}}$. Por lo tanto, tenemos

$$P(y|z, do(x)) = P(y|do(z)),$$

el cual ya ha sido calculado.

Tarea 3: Calcular $P(y|do(x))$ iii

Por lo tanto sustituyendo, nos queda:

$$
\begin{aligned}
P(y|do(x)) &= \sum_z P(y|z, do(x))P(z|do(x)) \\
&= \sum_z P(y|z, do(x))P(z|x) \\
&= \sum_z P(y|do(z), do(x))P(z|x) \\
&= \sum_z P(y|do(z))P(z|x) \\
&= \sum_z (\sum_x P(y|x, y)P(x))P(z|x) \\
&= \sum_z P(z|x) \sum_x P(y|x, z)P(x)
\end{aligned}
$$

y esto es igual a la fórmula de puerta delantera.

Tarea 4: Calcular $P(y, z|do(x))$

Tenemos

$$P(y, z|do(x)) = P(y|z, do(x))P(z|do(x)).$$

Los dos términos del lado derecho ya fueron derivados, por lo que tenemos

$$\begin{aligned} P(y, z|do(x)) &= P(y|do(z))P(z|do(x)) \\ &= P(z|x) \sum_{x} P(y|x, z)P(x) \end{aligned}$$

Tarea 5: Calcular $P(x, y|do(z))$

Tenemos

$$\begin{aligned} P(x, y|do(z)) &= P(y|x, do(z))P(x|do(z)) \\ &= P(y|x, z)P(x). \end{aligned}$$

El primer término se obtiene de la Regla 2 (bajo $G_{\underline{Z}}$) y el segundo por la Regla 3.

Inferencia causal por experimentos sustitutos (Surrogate experiments) i

- Supongamos que deseamos conocer el efecto causal de X en Y cuando $P(y|do(x))$ no es identificable y, por razones prácticas de costo o ética, no podemos controlar X mediante un experimento aleatorio.
- Surge la pregunta de si $P(y|do(x))$ se puede identificar al aleatorizar una variable de sustitución Z que es más fácil de controlar que X.
- Por ejemplo, si estamos interesados en evaluar el efecto de los niveles de colesterol (X) en la enfermedad cardíaca (Y), un experimento razonable para llevar a cabo sería controlar la dieta de los sujetos (Z), en lugar de ejercer un control directo sobre los niveles de colesterol en la sangre de los sujetos.

Inferencia causal por experimentos sustitutos (Surrogate experiments) ii

- Formalmente, este problema equivale a transformar $P(y|do(x))$ en expresiones en las que solo a los miembros de Z se les aplica el operador *do*. Usando el Teorema de las Reglas del cálculo do, se puede mostrar que las siguientes condiciones son suficientes para admitir una variable sustituta Z:
    1. X intercepta todos los caminos dirigidos de Z a Y; y
    2. $P(y|do(z))$ es identificable en $G_{\overline{Z}}$.
- Si la condicióń 1. se cumple, entonces podemos escribir $P(y|do(x)) = P(y|do(x), do(z))$, porque $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{XZ}}}$. Pero $P(y|do(x), do(z))$ representa el efecto causal de X sobre Y en un modelo gobernado por $G_{\overline{Z}}$, el cual - por la condición 2.- es identificable.

Inferencia causal por experimentos sustitutos (Surrogate experiments) iii

- Traducido al ejemplo del colesterol, estas condiciones requieren que no haya un efecto directo de la dieta sobre las condiciones cardíacas y que no haya factores de confusión entre los niveles de colesterol y las enfermedades cardíacas, a menos que podamos neutralizar dicha confusión mediante mediciones adicionales.
- Las Figuras 3(e) y (h) ilustran modelos en que ambas condiciones se mantienen. Con la figura (e), por ejemplo, obtenemos este estimando

$$P(y|do(x)) = P(y|x, do(z)) = \frac{P(y, x|do(x))}{P(x|do(z))}. \tag{1}$$

  Esto puede ser definido directamente primero aplicando la regla 3 para añadir $do(z)$

$$P(y|do(x)) = P(y|do(x), do(z))$$

Inferencia causal por experimentos sustitutos (Surrogate experiments) iv

porque $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}\overline{Z}}}$. Y luego aplicando la Regla 2 para cambiar $do(x)$ a x

$$P(y|do(x), do(z)) = P(y|x, do(z))$$

porque $(Y \perp\!\!\!\perp X|Z)_{G_{\underline{X}\overline{Z}}}$.

- De acuerdo con (1), sólo un nivel de Z es suficiente para la identificación de $P(y|do(x))$ para cualesquiera valores de y y x. En otras palabras, Z no necesita ser variado en absoluto; puede mantenerse constante por medios externos y, si las suposiciones incorporadas en G son válidas, el lado derecho de (1) debe alcanzar el mismo valor independientemente del nivel (constante) en el que Z se mantiene. Sin embargo, en la práctica, se necesitarán varios niveles de Z para asegurar que se obtengan suficientes muestras para cada valor deseado de X.

Inferencia causal por experimentos sustitutos (Surrogate experiments) V

- Por ejemplo, si estamos interesados en la diferencia $E(Y|do(x)) - E(Y|do(x'))$, donde x y x' son dos niveles de tratamiento, entonces deberíamos elegir dos valores z y z' de Z que maximiza el número de muestras en x y x' (respectivamente) y luego estimar

$$E(Y|do(x)) - E(Y|do(x')) = E(Y|x, do(z)) - E(Y|x', do(z')).$$

Inferencia causal por experimentos sustitutos (Surrogate experiments) vi

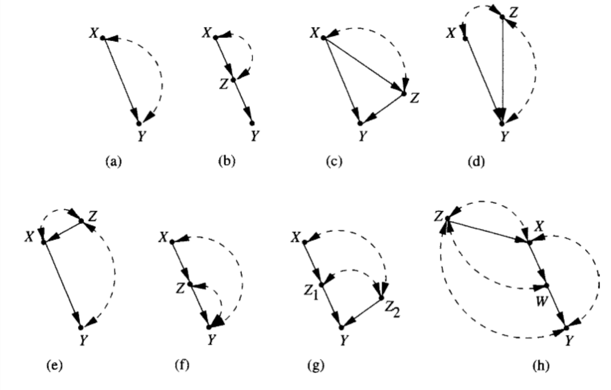


Figura 3: Modelos en los cuales $P(y|do(x))$ no es identificable.