# Redes Bayesianas Causales

## Modelos Gráficos Causales

Sebastián Bejos y Claudio López

16 de julio de 2019

Instituto Nacional de Astrofísica Óptica y Electrónica

# Introducción

Las ventajas de construir modelos (basados en un representación por medio de DAGs) sobre información causal por sobre información asociativa son varias:

1. Los juicios para la construcción del modelo son más significativos, más accesibles y, por lo tanto, más confiables.
2. Permiten representar y responder a cambios externos o espontáneos. Cualquier reconfiguración local de los mecanismos en el entorno se puede traducir, con solo una pequeña modificación, en una reconfiguración isomórfica de la topología de la red (DAG). (Agentes deliberativos vs reactivos)

# Redes Causales como Oráculos para Intervenciones

- Esta flexibilidad se basa en el supuesto de que cada relación padre-hijo en la red representa un mecanismo físico estable y autónomo, en otras palabras, que es posible cambiar una de esas relaciones sin cambiar las demás.
- Organizar el conocimiento con estas configuraciones modulares, permite predecir el efecto de las intervenciones externas con un mínimo de información adicional.

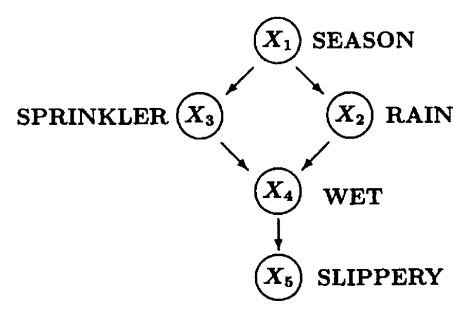
# Redes Causales como Oráculos para Intervenciones

- Una distribución conjunta nos dice que tan probable son los eventos y como cambiarán las probabilidades con observaciones posteriores.
- Un modelo causal, además nos dice cómo cambiarían estas probabilidades como resultado de intervenciones externas.

# Redes Causales como Oráculos para Intervenciones

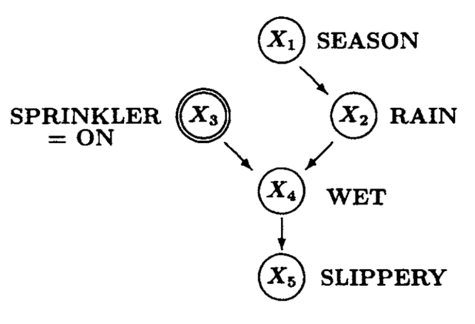
- En lugar de especificar una nueva función de probabilidad para cada una de las muchas intervenciones posibles.
- El efecto general de una intervención se puede predecir modificando la factorización de la distribución conjunta y utilizando el producto modificado se calcula una nueva función de probabilidad.

# Redes Causales como Oráculos para Intervenciones




$$P(x_1, x_2, x_3, x_4, x_5) = P(x_1)P(x_2 \mid x_1)P(x_3 \mid x_1)P(x_4 \mid x_2, x_3)P(x_5 \mid x_4)$$

# Redes Causales como Oráculos para Intervenciones




$$P_{X_3=On}(x_1, x_2, x_4, x_5) = P(x_1)P(x_2 \mid x_1)P(x_4 \mid x_2, X_3 = On)P(x_5 \mid x_4)$$

# Redes Bayesianas Causales

**Definición**

*Sea $P(\mathbf{v})$ una distribución de probabilidad sobre un conjunto de variables $\mathbf{V}$, y sea $P_{\mathbf{x}}(\mathbf{v})$ la distribución resultante de la intervención $do(\mathbf{X} = \mathbf{x})$ que establece un subconjunto de variables $\mathbf{X}$ como contantes $\mathbf{x}$. Denotemos por $\mathbf{P}*$ al conjunto de todas las distribuciones intervencionales $P_{\mathbf{x}}(\mathbf{v})$, $\mathbf{X} \subseteq \mathbf{V}$, incluyendo a $P(\mathbf{v})$, que representa ninguna la intervención (i.e., $\mathbf{X} = \emptyset$). Una DAG $G$ es una* **Red Bayesina Causal** *compatible con $\mathbf{P}*$ si y solo si, se cumplen la siguientes tres condiciones para toda $P_{\mathbf{X}} \in \mathbf{P}*$:*

(i) $P_{\mathbf{X}}(\mathbf{v})$ *es Markov relativa a $G$.*

(ii) $P_{\mathbf{X}}(\mathbf{v}_i) = 1$ *para toda $\mathbf{V}_i \in \mathbf{X}$ cada vez que $\mathbf{v}_i$ es consistente con $\mathbf{X} = \mathbf{x}$.*

(iii) $P_{\mathbf{x}}(\mathbf{v}_i \mid \mathbf{pa}_i) = P(\mathbf{v}_i \mid \mathbf{pa}_i)$ *para toda $\mathbf{V}_i \notin \mathbf{X}$ cada vez que $\mathbf{pa}_i$ sea consistente con $\mathbf{X} = \mathbf{x}$, i.e., cada $P(\mathbf{v}_i \mid \mathbf{pa}_i)$ permanezca invariante a intervenciones que no involucren a $\mathbf{V}_i$.*

- La definición anterior, impone restricciones en el espacio intervencionista $\mathbf{P}*$ que nos permite codificar económicamente este vasto espacio, en forma de una única red bayesiana $G$.
- Estas restricciones nos permiten calcular la distribución $P_X(v)$ resultante de cualquier intervención $do(X = x)$ como una factorización truncada.

$$P_X(v) = \prod_{\{i | V_i \notin X\}} P(v_i \mid pa_i), \text{para toda } v \text{ consistente con } x,$$

la cual se sigue (e implica) de las condiciones (i)-(iii).

# Redes Bayesianas Causales

Si $G$ es una red bayesiana causal con respecto a $P*$ se tienen la siguiente propiedades:

- **Propiedad 1** Para toda $i$, $P(v_i \mid pa_i) = P_{pa_i}(v_i)$.
  Esta propiedad hace que cada conjunto de padres $PA_i$ sea exógeno en relación con su hijo $V_i$, asegurando que la probabilidad condicional $P(v_i \mid pa_i)$ coincida con el efecto (sobre $V_i$) de establecer $PA_i$ en $pa_i$ por control externo.
- **Propiedad 2** Para toda $i$ y para todo sunconjunto $S$ de variables, disjunto de $\{V_i, PA_i\}$, se tiene $P_{pa_i,S}(v_i) = P_{pa_i(v_i)}$.
  Esta propiedad expresa la noción de invariancia; una vez que controlamos sus causas directas $PA_i$, ninguna otra intervención afectará la probabilidad de $V_i$.

# 1.4 Modelos Causales Funcionales

El enfoque tomado bajo la concepción cuasi-determinista de la causalidad de Laplace.

- En primer lugar, la concepción laplaciana es más general. Cada modelo estocástico puede ser emulado por muchas relaciones funcionales (con entradas estocásticas), pero no al revés; Las relaciones funcionales solo pueden ser aproximadas, como un caso limitante, utilizando modelos estocásticos.
- Segundo, la concepción laplaciana está más en sintonía con la intuición humana. Los pocos experimentos mecánicos cuánticos esotéricos que entran en conflicto con las predicciones de la concepción laplaciana evocan sorpresa e incredulidad y exigen que los físicos renuncien a intuiciones profundamente arraigadas sobre la localidad y la causalidad (Maudlin, 1994).

## 1.4.1 Ecuaciones Estructurales

$$x_i = f_i(pa_i, u_i), i = 1, ..., n \tag{1}$$

$$x_i = \sum_{k \neq i} \alpha_{ik} x_k + u_i, i = 1, ..., n \tag{2}$$

Ejemplo: Modelo econométrico canónico que relaciona el precio y la demanda a través de las ecuaciones.

$$q = b_1 p + d_1 i + u_1 \tag{3}$$

$$p = b_2 q + d_2 w + u_2 \tag{4}$$

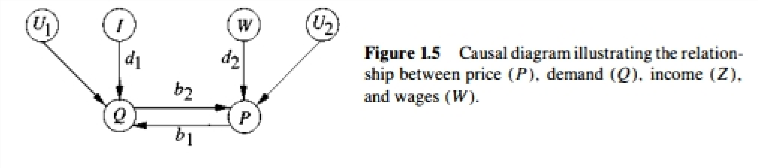



**Figura 1:** Representación causal de 3 y 4

## 1.4.1 Ecuaciones Estructurales

Para el ejemplo 1.2 mostrado anteriormente, su representación en SEM es:

$$x_1 = u_1 \tag{5}$$

$$x_2 = f_2(x_1, u_2) \tag{6}$$

$$x_3 = f_3(x_1, u_3) \tag{7}$$

$$x_4 = f_4(x_3, x_2, u_4) \tag{8}$$

$$x_5 = f_5(x_4, u_5) \tag{9}$$

Ahora bien, las funciones se ven representadas de las siguiente forma:

$$x_2 = [(X_1 = winter)v(X_1 = fall)vu_2] \wedge u_2' \tag{10}$$

$$x_3 = [(X_1 = summer)v(X_1 = spring)vu_3] \wedge u_3' \tag{11}$$

$$x_4 = [x_2 \vee x_3 \vee u_4] \wedge u_4' \tag{12}$$

$$x_5 = (x_4 \vee u_5) \wedge u_5' \tag{13}$$

- Predicciones (por ejemplo, ¿el pavimento sería resbaladizo si encontramos el aspersor apagado?).
- Intervenciones (por ejemplo, ¿el pavimento sería resbaladizo si nos aseguramos de que el aspersor esté apagado?).
- Contractual (por ejemplo, ¿el pavimento estaría resbaladizo si el aspersor hubiera estado apagado, dado que el pavimento no es resbaladizo y el rociador está encendido?).

## 1.4.2 Predicciones probabilísticas en modelos causales

Por medio del uso de los modelos bayesianos, donde ya esta comprobado su uso en predicciones de eventos, sobre los modelos causales debido a su condición markoviana, se puede predecir acontecimientos causales al estipular las condiciones iniciales.

Hay una ventaja adicional de basar los modelos de predicción en mecanismos causales que se derivan de consideraciones de estabilidad. Cuando algunas condiciones en el entorno experimentan un cambio, generalmente son solo unos pocos mecanismos causales los que se ven afectados por el cambio; el resto permanece inalterado. Es más sencillo entonces reevaluar (juzgar) o reestimar (estadísticamente) los parámetros del modelo sabiendo que el cambio simbólico correspondiente también es local, que involucra solo unos pocos parámetros, que reestimar todo el modelo desde cero.

## 1.4.3 Intervenciones y efectos causales en modelos funcionales.

La representación de intervenciones del modelo funcional ofrece mayor flexibilidad y generalidad que la de un modelo estocástico.

**Primero** , el análisis de las intervenciones puede extenderse a los modelos cíclicos, como el de la Figura 1, para responder a preguntas relacionadas con las políticas (por ejemplo, ¿cuál sería la cantidad de demanda si controlamos el precio en P0?).

**Segundo**, las intervenciones que involucran la modificación de los parámetros de las ecuaciones (como $b_1$ y $d_1$ en (3)) se comprenden más fácilmente que los descritos como modificadores de las probabilidades condicionales, tal vez debido a una estabilidad física Los mecanismos normalmente se asocian con ecuaciones y no con probabilidades condicionales.

**Tercero**, el análisis de los efectos causales en modelos no markovianos se simplificará en gran medida utilizando modelos funcionales. La razón es: hay infinitas probabilidades condicionales $P(X_i \setminus pa_i)$ pero solo un número finito de funciones $x_i = f_i(pa_i, u_i)$ entre las variables discretas $X_i$

# 1.4.4 Contrafactuales en modelos funcionales

Ciertas oraciones contrafactuales, como comentamos anteriormente, no pueden definirse en el marco de las redes causales estocásticas. Para ver las dificultades, consideremos la red bayesiana causal más simple posible que consiste en un par de variables binarias independientes (y, por lo tanto, no conectadas) X e Y. Una red de este tipo surge, por ejemplo, en un ensayo clínico controlado (es decir, aleatorio) cuando encontramos que un tratamiento X no tiene efecto en la distribución de la respuesta Y de los sujetos, que puede significar recuperación ($Y = 0$) o muerte ($Y = 1$). Supongamos que un sujeto dado, Joe, ha tomado el tratamiento y murió; preguntamos si la muerte de Joe se produjo debido al tratamiento, a pesar del tratamiento o independientemente del tratamiento. En otras palabras, pedimos la probabilidad Q de que Joe hubiera muerto si no hubiera sido tratado.

## 1.4.4 Contrafactuales en modelos funcionales

Supongamos el caso mas extremo al aplicar un tratamiento para cierta enfermedad: el 50 % de los pacientes se recuperan y 50 % mueren tanto con o sin el tratamiento.

P(Y/X) = 1/2 para todo los valores x y y.

Preguntando qué porcentaje de sujetos Q que murieron bajo tratamiento se hubieran recuperado si no hubieran tomado el tratamiento.

Teniendo en cuenta que existen causas que no podemos observar y que impiden explicar los resultados se generan los grafos.

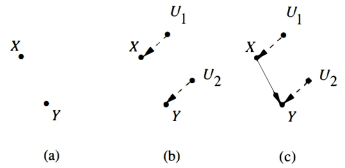



**Figure 1.6** (a) A causal Bayesian network that represents the distribution of (1.47). (b) A causal diagram representing the process generating the distribution in (a), according to model 1. (c) Same, according to model 2. (Both $U_1$ and $U_2$ are unobserved.)

**Figura 2**

## 1.4.4 Contrafactuales en modelos funcionales

Proporcionamos un análisis de los enfoques del mundo más cercano y del resultado potencial y los comparamos con el enfoque del modelo estructural, que se describirá a continuación, en el cual los contrafactuales se derivan de (y de hecho se definen por) un modelo causal funcional.

Suponemos la misma distribución de probabilidad con dos variables independientes pero con una probabilidad conjunta.

$P(y/x)=0.25$ para todo x y y.

Ahora presentamos dos modelos funcionales, cada uno de los cuales genera la probabilidad conjunta de , pero cada uno de ellos da un valor diferente a la cantidad de interés, $Q =$ la probabilidad de que un sujeto que falleció bajo tratamiento ($x = 1$, $y = 1$) tenga recuperado ($y = 0$) si él o ella no hubiera sido tratado ($x = 0$).

## 1.4.4 Contrafactuales en modelos funcionales

Se proponen 2 modelos:

Modelo 1 Figura 2.b:

$$x = u_1 \tag{14}$$

$$y = u_2 \tag{15}$$

Modelo 2 Figura 2.c:

$$x = u_1 \tag{16}$$

$$y = x * u_2 + (1 - x)(1 - u_2) \tag{17}$$

Donde $U_1$ y $U_2$ son variables binarias independientes con $P(U_1=1)=P(U_2=1)=1/2$.

## 1.4.4 Contrafactuales en modelos funcionales

| **Model 1** | $u_2 = 0$ | | $u_2 = 1$ | | Marginal | |
|---|---|---|---|---|---|---|
| | $x = 1$ | $x = 0$ | $x = 1$ | $x = 0$ | $x = 1$ | $x = 0$ |
| $y = 1$ (death) | 0 | 0 | 0.25 | 0.25 | 0.25 | 0.25 |
| $y = 0$ (recovery) | 0.25 | 0.25 | 0 | 0 | 0.25 | 0.25 |

| **Model 2** | $u_2 = 0$ | | $u_2 = 1$ | | Marginal | |
|---|---|---|---|---|---|---|
| | $x = 1$ | $x = 0$ | $x = 1$ | $x = 0$ | $x = 1$ | $x = 0$ |
| $y = 1$ (death) | 0 | 0.25 | 0.25 | 0 | 0.25 | 0.25 |
| $y = 0$ (recovery) | 0.25 | 0 | 0 | 0.25 | 0.25 | 0.25 |

**Figura 3:** Tablas de contingencia que muestran las distribuciones $P(x,y,u_2)$ y $P(x,y)$ para los dos modelos discutidos en el texto.

El modelo I corresponde al tratamiento (X) que no tiene efecto en ninguno de los sujetos; En el modelo 2, todos los sujetos se ven afectados por el tratamiento. La razón por la que los dos modelos producen la misma distribución es que el modelo 2 describe una mezcla de dos subpoblaciones. En uno ($U_2 = 1$), cada sujeto muere ($y=1$) si y solo si se trata; en el otro ($U_2=0$), cada sujeto se recupera ($y = 0$) si y solo si se trata. Las distribuciones $P(x,y,U_2)$ y $P(x,y)$ correspondientes a estos dos modelos se muestran en las tablas de la Figura 3.

# 1.4.4 Contrafactuales en modelos funcionales

El valor de Q difiere en estos dos modelos. En el modelo 1, Q se evalúa a cero, porque los sujetos que murieron corresponden a $U_2=1$ y, dado que el tratamiento no tiene efecto en y, el cambio de X de 1 a 0 todavía daría $y = 1$. En el modelo 2, sin embargo, Q evalúa la unidad, porque los sujetos que murieron bajo tratamiento deben corresponder a $U_2=1$ (es decir, aquellos que mueren si son tratados), lo que significa que se recuperarán si y solo si no se tratan.

La primera lección de este ejemplo es que los modelos causales estocásticos son insuficientes para calcular las probabilidades de contrafactuales; el conocimiento del proceso real detrás de $P(y/x)$ es necesario para el cálculo.24 Una segunda lección es que un modelo causal funcional constituye un objeto matemático suficiente para el cálculo (y definición) de tales probabilidades.

## 1.4.4 Contrafactuales en modelos funcionales

Se pueden generalizar a cualquier modelo causal M, con 3 pasos de la siguiente manera. Dada la evidencia e, para calcular la probabilidad de Y=y bajo la condición hipotética X=x (donde X es un subconjunto de variables), aplique los siguientes tres pasos a M.

1. **(abducción:** Actualice la probabilidad P(u) para obtener P(u/e).
2. **(acción):** Reemplace las ecuaciones correspondientes a las variables en el conjunto X por las ecuaciones X=x.
3. **(predicción):** use el modelo modificado para calcular la probabilidad de Y=y.

En las metáforas temporales, este procedimiento de tres pasos se puede interpretar de la siguiente manera. El paso 1 explica el pasado (U) a la luz de la evidencia actual e; el paso 2 dobla el curso de la historia (como mínimo) para cumplir con la condición hipotética X=x; el paso 3 predice el futuro (Y) basado en nuestra nueva comprensión del pasado y nuestra condición recién establecida, X=x.

## 1.4.4 Contrafactuales en modelos funcionales

Recordando que para cada valor u de U hay una solución única para Y, está claro que el paso 3 siempre da una solución única para la probabilidad necesaria; simplemente resumimos las probabilidades P(u/e) asignadas a todos aquellos u que producen Y=y como solución.

El modelo de tres pasos del razonamiento contrafactual también descubre la razón real por la que los modelos causales estocásticos son insuficientes para calcular las probabilidades de los contrafactuales. Debido a que las variables U no aparecen explícitamente en los modelos estocásticos, no podemos aplicar el paso 1 para actualizar P(u) con la evidencia a la mano. Esto Para estos, debemos hacer algunas suposiciones sobre la forma de las funciones $f_i$ y las probabilidades de los términos de error. (Cap 7 y 9).

## 1.4.4 Contrafactuales en modelos funcionales

Las consideraciones anteriores implican además que las tres tareas enumeradas en el principio de esta sección (predicción, intervención y contrafactuales) forman un jerárquico natural de tareas de razonamiento causal, con niveles crecientes de refinamiento y demanda creciente: sobre el conocimiento requerido para realizar estas tareas. **La predicción** es la más simple de las tres, y solo requiere una especificación de una función de distribución conjunta. **El análisis de las intervenciones** requiere una estructura causal además de una distribución conjunta. Finalmente, procesar los **contrafactuales** es la tarea más difícil porque requiere cierta información sobre las relaciones funcionales y/o la distribución de los factores omitidos.