《因果论》第二章汇报：因果关系推断理论

梁克维　张逸凡

浙江大学数学科学学院

2024 年 10 月 13 日

基本直觉

人的直觉常常被认为是“不科学的”. 然而，这种直觉在因果关系的发现中却扮演着重要的角色. 因为人们在现实生活中，往往能够通过自己的直觉构造出合理的因果逻辑，因此我们试图找到模拟这种感知的计算模型.

人的基本直觉往往与以下两个问题有关:

- 时序关系是否是发现因果关系的充分条件，或必要条件?
- 控制变量法是否是发现因果关系的本质?

因果发现框架

为以后讨论的方便，我们先给出因果结构和因果模型的严格定义.

[定义](因果结构)
一组变量 V 的因果结构是一个有向无环图 (DAG), 其中每个节点对应 V 中的一个变量，每个链接表示对应变量之间的直接函数关系.

[定义](因果模型)
因果模型是由因果结构 D 及与 D 相容的参数集 Θ_D 组成的 $M =< D, \Theta_D >$. 参数集 Θ_D 为每个变量 $X_i \in V$ 赋予函数 $x_i = f_i(pa_i, u_i)$，并为每个 u_i 赋予概率值 $P(u_i)$，其中 PA_i 是 D 中 X_i 的父代变量，U_i 是根据 $P(U_i)$ 分布的随机扰动，所有的 U_i 相互独立.

- 可以看出，作者认为在大自然中，互相关联的物体之间具有函数关系，而随机性则由随机变量 U_i 引入.
- 在如上的定义中，因果结构是有向无环的以及所有的 U_i 相互独立是作者做出的假设. 后者是为了保证因果模型具有马尔可夫性，即以 D 中的父代变量为条件时，每个变量都独立于它的所有非后代变量.
- 假如我们知道的信息足够多 (以至于我们了解事物的微观细节)，那么马尔可夫性自然成立. 然而，当我们进行宏观抽象时，必然会丢失一些性质，显然我们的前辈认为马尔可夫性是抽象过程中应当保留的性质.

若我们观察到的模型不满足马尔可夫性，那只能解释为我们观察到的变量不够多. 假如我们承认“潜”变量的存在，那么马尔可夫性仍然成立. 为此，我们引入潜在结构的定义:

[定义](潜在结构)
潜在结构定义为 $L =< D, O >$，其中 D 是 V 上的因果结构，而 $O \subseteq V$ 是一个观测变量集.

本章主要研究的问题是: 科学家检测到“观察变量”的子集 $O \subseteq V$ 的概率分布 $P_{[o]}$ 后从中推断出变量之间因果关系的方法. 为此，我们需要引入两个假设.

假设 1：奥卡姆剃刀原则

- 从哲学角度来讲，奥卡姆剃刀原则表达了对简单的偏好，即不引入不必要的假设.
- 从实际应用角度来讲，奥卡姆剃刀原则体现在防止过拟合.

以此原则为指引，我们可以给出结构偏好以及极小性的定义.

[定义](结构偏好)
潜在结构 $L =< D, O >$ 优于潜在结构 $L' =< D', O >$(记为 $L \preceq L'$) 当且仅当对于每个 Θ_D, 存在 $\Theta'_{D'}$ 使得 $P_{[O]}(< D', \Theta'_{D'} >) = P_{[O]}(< D, \Theta_D >)$. 两个潜在结构是等价的 (记为 $L' \equiv L$)，当且仅当 $L \preceq L'$ 且 $L' \preceq L$.

[定义](极小性)
潜在结构 L 相对于潜在结构类 $\mathcal{L}$ 是极小的，当且仅当 $\mathcal{L}$ 中没有结构严格优于 L.

当然，我们所构建的因果结构也不能只顾着简单，也需要与观察到的分布一致，即：

[定义](一致性)
潜在结构 $L =< D, O >$ 在 O 上的分布与 $\hat{P}$ 一致，如果存在参数化 Θ_D 使得 $P_{[O]}(< D, \Theta_D >) = \hat{P}$.

有了上述定义后，我们终于可以给出因果推断的定义：

[定义](因果推断)
给定 $\hat{P}$，当且仅当在每个与 $\hat{P}$ 一致的极小潜在结构中存在从 C 到 E 的有向路径，则变量 C 对变量 E 有因果影响.

假设 2：稳定分布

稳定性原则是指，假设 P 中包含的所有独立性都是稳定的，即它们由模型 D 的结构所限定，对参数 Θ_D 的任何变化都保持不变. 根据这种思想，我们引出如下定义：

[定义](稳定性)
令 $I(P)$ 表示 P 中蕴含的所有条件独立关系的集合. 当且仅当对于任意的参数 Θ'_D,$I(P(<D,\Theta_D>)) \subseteq I(P(<D,\Theta'_D>))$, 因果模型 $M=<D,\Theta_D>$ 生成稳定的分布.

稳定性条件表明，当我们将参数从 Θ_D 变化到 Θ'_D 时，P 中的独立性均不会被破坏，即我们观察到的条件独立并不是”碰巧”观察到的，而是源于大自然稳定的机制. 这种假设的依据是，参数乘积之间的严格相等在任何概率空间中的勒贝格测度为零.

获取 DAG 结构

在稳定性假定之下，只要没有隐藏变量，每个分布都有唯一的极小因果结构. 但是此处的”唯一”是在 DAG 观察等价意义下的. 因此，我们所能做的就是找到一个 DAG 观察等价类的图模型表示. 为此，我们引入模式的定义：

[定义](模式)
模式是一些边有向、一些边无向的图，其中有向边表示 DAG 观察等价类中的每个结构所共有的箭头，而无向边代表其中可能的矛盾：在某些等价结构中，它们指向一种方向，而在另一些结构中，它们指向另一种方向.

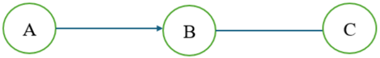
基于此，Verma 和 Pearl(1990) 提出了以概率分布作为输入，输出一个 DAG 等价类的模式算法.

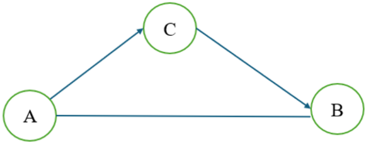
IC 算法 (Inductive Causation)

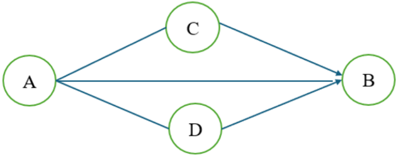
输入：变量集 V 上的稳定分布 $\hat{P}$
输出：与 $\hat{P}$ 相容的模式 $H(\hat{P})$

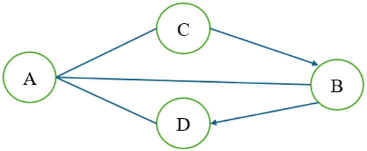
- (1). 对于 V 中的每对变量 a 和 b，寻找极小集合 S_{ab} 使得 $(a \perp\!\!\!\perp b|S_{ab})$ 在 $\hat{P}$ 中成立. 构建无向图 G 使得当且仅当 S_{ab} 不存在时，连接节点 a 和 b.
- (2). 对于每对具有公共邻居 c 的非邻接变量 a 和 b，检查 $c \in S_{ab}$ 是否成立. 若成立，则继续；若不成立，则添加从 a 和 b 指向 c 的有向边.
- (3). 根据以下两个条件，为尽可能多的无向边定向：(i) 任何可选的方向会产生一个新的头对头结构；(ii). 任何可选的方向都会产生一个有向环.

对于上述算法第三步的几个例子









IC 算法没有详细说明步骤 1 和步骤 3，一些研究者提出了改进以优化这两个步骤.

- 一些研究者已经找到在多项式时间复杂度内完成步骤 1 的方法.
- Meek(1995) 指出，若反复应用上述 4 个例子所对应的条件，足以完成 IC 算法的步骤 3.
- Dor 和 Tarsi(1992) 提出了一种算法，在多项式时间内检验是否能将 IC 算法第 2 步生成的部分定向的无环图通过步骤 3 变成完全定向的无环图.

重组潜在结构

上述的 IC 算法仅仅在科学家能够观察到所有变量时成立，然而事实往往并非如此. 因此，我们需要寻找出当大自然" 隐藏" 了一些变量时，从观察到的分布 $\hat{P}$ 推断出因果关系的方法. 为此，我们需先定义潜在结构 $L =< D, O >$ 在观察变量集 O 上的投影.

[定义](投影)
潜在结构 $L_{[O]} =< D_{[O]}, O >$ 是潜在结构 L 的投影，当且仅当
(1).$D_{[O]}$ 中的每个未观测变量都恰好是两个非邻近的可观测变量的共同原因，且该变量无父代变量.
(2). 对于 L 生成的每个稳定分布 P，都存在一个 $L_{[O]}$ 生成的稳定分布 P'，使得 $I(P_{[O]}) = I(P'_{[O]})$.

Verma(1993) 证明了如下定理:

[定理](Verma 1993)
任何潜在结构都至少有一个在其观察集 O 上的投影.

这个定理表明，可以很方便的用一个双向图来表示投影，途中仅将可观测变量作为节点 (即不显示隐藏变量). 图中的每个双向链接代表了对应于该链接两端变量的共同隐藏原因.

有了以上准备工作，我们可以对 IC 算法加以修改，称之为 IC* 算法. 该算法以稳定分布 $\hat{P}$ 为输入，输出一个标记的模式. 该模式的边具有以下四种类型，并且是一个部分有向无环图.

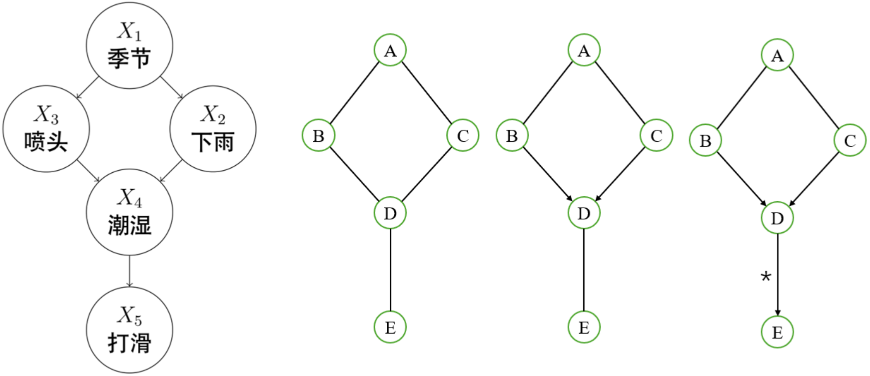
- 1. 有标记箭头 $a \overset{*}{\rightarrow} b$ 表示前在模型中 a 到 b 的有向路径.
- 2. 无标记箭头 $a \rightarrow b$ 表示前在模型中从 a 到 b 的有向路径或者一个潜在的公共原因 $a \leftarrow L \rightarrow b$.
- 3. 双向边 $a \leftrightarrow b$ 表示潜在模型中的一个潜在公共原因 $a \leftarrow L \rightarrow b$.
- 4. 无向边 $a - b$ 表示潜在模型中 $a \leftarrow b$，或者 $a \rightarrow b$，或者 $a \leftarrow L \rightarrow b$.

IC* 算法 (Inductive Causation with Latent Variables)

输入：稳定分布 $\hat{P}$(对应某种潜在结构)
输出：一个被标记的模式 $core(\hat{P})$

- (1). 对于 V 中的每对变量 a 和 b，寻找极小集合 S_{ab} 使得 $(a \perp\!\!\!\perp b|S_{ab})$ 在 $\hat{P}$ 中成立. 构建无向图 G 使得当且仅当 S_{ab} 不存在时，连接节点 a 和 b.
- (2). 对于每对具有公共邻居 c 的非邻接变量 a 和 b，检查 $c \in S_{ab}$ 是否成立. 若成立，则继续；若不成立，则添加从 a 和 b 指向 c 的有向边.
- (3). 根据以下两个条件，(递归的) 添加尽可能多的箭头，并标记尽可能多的边：
  (i) 对于每对具有公共邻居 c 的非相邻变量 a 和 b，如果 a 和 c 之间的链接有箭头指向 c，而 c 和 b 之间的链接没有箭头指向 c，则在 c 和 b 之间的链接上添加箭头指向 b，并标记这条链接，得到 $c \xrightarrow{*} b$.
  (ii). 如果 a 和 b 相邻，并且由一条从 a 到 b 的 (严格由标记的链接组成的) 有向路径，则在 a 和 b 之间的链接上添加一个指向 b 的箭头.

一个实例



a) b) c) d)

因果关系推断的局部准则

在 IC* 算法的输出中，我们称有标记单向的链接为真实原因，无标记单向的链接为潜在原因，双向的链接为虚假关联. 产生这些标签的条件可以用作各种因果关系的定义.

[定义](潜在原因)
如果下列条件成立，则变量 X 对变量 Y 有潜在的因果影响:

- 1.X 和 Y 在每个情形下均相关.
- 2. 存在变量 Z 和情形 S，使得:
  (1). 给定 S 条件下，X 和 Z 独立;
  (2). 给定 S 条件下，Z 和 Y 相关.

[定义](真实原因)

如果存在一个变量 Z 满足以下二者之一，则变量 X 对变量 Y 有真实的因果影响：

- 1.X 和 Y 在任何情形中均相关，且存在情形 S 满足
  (1).Z 是 X 的潜在原因;
  (2). 给定 S 时，Z 和 Y 相关;
  (3). 给定 $S \cup X$ 时，Z 和 Y 独立.
- 2. X 和 Y 在准则 1 中定义的关系传递闭包中.

[定义](虚假关联)
两个变量 X 和 Y 是虚假关联的，如果它们在某些情形下相关，且存在变量 Z_1, Z_2 以及情形 S_1, S_2 使得：

- 给定 S_1， Z_1 和 X 相关;
- 给定 S_1， Z_1 和 Y 独立;
- 给定 S_2， Z_2 和 Y 相关;
- 给定 S_2， Z_2 和 Y 独立.

当时间信息可用时，对于真实原因和虚假关联的定义可大大简化，因为与 X 相邻且在 X 之前的每个变量都可作为 X 的“潜在原因”. 此外，只要情形 S 被限制在 X 之前，则不再要求邻接.

[定义](具有时间信息的真实因果)
变量 X 对 Y 有因果影响，如果存在第三个变量 Z 和情形 S，二者都在 X 之前出现，且满足：

- 1. 给定 S 时，Z 和 Y 相关;
- 2. 给定 $S \cup X$ 时，Z 和 Y 独立.

[定义](具有时间信息的虚假关联)
两个变量 X 和 Y 是虚假关联的，如果它们在某些情形 S 中相关，X 先于 Y，且存在变量 Z 满足：

- 1. 给定 S，Z 和 Y 独立；
- 2. 给定 S，Z 和 X 相关.

我们发现，所有这些变量之间的因果关系的定义中，两个变量 $(X$ 和 $Y)$ 之间的因果关系的判断准则都需要第三个变量 Z 来确定. 这并不奇怪，因为因果断言的本质是规定 X 和 Y 在第三个变量影响下的行为，即 "没有操纵，就没有因果". 因此，IC 算法和 IC* 算法本质上是试图识别出大自然对变量集进行的控制变量实验.

非时间因果与统计时间

从非时间数据中确定因果影响方向的问题引发了一些关于时间与因果解释之间有趣的哲学问题，即：从非时间数据中确定的因果方向一定与物理时间一致吗？为更好的描述这一问题，我们先定义统计时间：

[定义](统计时间)
给定一个经验分布 P，P 的统计时间是满足以下条件的变量的任意排序，即与至少一个与 P 相容的极小因果结构一致.

直觉使人们猜测，在大多数自然现象中，物理时间与至少一个统计时间一致. 然而，这一点目前还没有得到明确的证明.