因果发现算法的实现

编程实践手册

集智俱乐部因果读书会

计算技术研究所 李奉治

2021 年 5 月 29 日

目录

1 SGS-Algorithm

SGS 算法由 Peter Spirtes、Clark Glymour 和 Richard Scheines 三人与 1989 年提出 [4]，名字取自三人姓氏首字母。原文并未以伪代码形式展示其运行流程，后在 1991 年的文章 [3] 中提及 SGS 算法的正确性可以由 Verma 的 1990 年文章所证明 [6]，并在文中将 SGS 算法被描述为

`Input:` A distribution P or samples obtained from P.

`Output:` A collection of directed acyclic graphs

1. Strat with the complete undirected graph.
2. For each vertex pair (a, b), remove the undirected edge between a and b is and only if $I(a, S, b)$ for some subsets not containing a or b. Call this undirected graph G.
3. For each triple (a, b, c) of vertices such that a and b are adjacent in G, b and c are adjacent in G, and a and c are not adjacent in G, direct the edges $a - b$ and $b - c$ into b if and only if for every set S of vertices containing b but not a or c, $I(a, S, c)$.
4. Output all orientations of the graph consistent with (2)

2 IC-Algorithm

IC(Inductive Causation) 算法由 TS Verma 和 Judea Pearl 于 1990 年提出 [6]，结合 Jueal Pearl 在其 *Causality* 一书的 2.5 节中的算法描述 [2]，其算法的核心思想为：

`Input`：$\hat{P}$ a stable distribution on a set V of variables.

`Output`：A Pattern $H(\hat{P})$ compatible with $\hat{P}$.

1. For each pair of variables a and b in V, search for a set S_{ab} such that a and b are independent in $\hat{P}$ conditioned on S_{ab}. If there is no such S_{ab}, place an undirected link between the variables.
2. For each pair of non-adjacent variables a and b with a common neighbor c, check if $c \in S_{ab}$
   - If it is, then continue.
   - If it is not, then add arrowheads pointing at c, (i.e. $a \to c \leftarrow b$).
3. In the partially directed graph that results, orient as many of the undirected edges as possible subject to two conditions:
   (a) the orientation should not create a new v-structure;
   (b) the orientation should not create a directed cycle.

定义 1 ***V*-结构 *(V-structure)*** 对于任意三个结点 A、B 和 C，如果满足结构 $A \to B \leftarrow C$ 且结点 A 和 C 之间不相邻，则称此为一个 V-结构。

相较于 SGS 算法，IC 算法的增加了第 3 步的内容，继续为 PDAG(Partially DAG) 中的无向边确认方向，并最终输出一个 PDAG。而 SGS 算法输出的为所有可能的 DAG。

Judea Pearl 在 *Causality* 中指出，IC 算法提出时第 1 步和第 3 步仅仅是描述性的，未给出具体的方法。第 1 步后来由 PC 算法给出的高效的实现。

实际上，在 IC 算法提出的原文中 [6]，第 3 步也提供了以下的方法

1. Form core($\hat{P}$) by recursively adding arrowheads according to the following two rules:
   - If $a - b$ and there is a strictly directed path from a to b then add an arrowhead at b.
   - If a and b are not adjacent but $a \to c$ and $c - b$, then direct the link $c \to b$.
2. If $a \to b$ then mark every uni-directed link $b \to c$ in which c is not adjacent to a.

但并没有证明此方法的充分性。

后在 1992 年的一篇文章中，TS Verma 给出了更详细的四条规则 [7]，这四条规则在 *Causality* 一书中表述如下：

- R1: Orient $b - c$ into $b \to c$ whenever there is an arrow $a \to b$ such that a and c are nonadjacent.
- R2: Orient $a - b$ into $a \to b$ whenever there is chain $a \to c \to b$.
- R3: Orient $a - b$ into $a \to b$ whenever there are two chains $a - c \to b$ and $a - d \to b$ such that c and d are nonadjacent.
- R4: Orient $a - b$ into $a \to b$ whenever there are two chains $a - c \to d$ and $c \to d \to b$ such that c and b are nonadjacent and a and d are adjacent.

实际上，如果是基于一个已经标定好 V-结构的 PDAG，只需要应用前三条规则即可。

Meek 在 1995 年证明了这四条规则的充分性 [1]。上述四条规则的必要性证明如下：

- R1: 反之如果 $b \leftarrow c$，则会形成新的 V-结构 $a \to b \leftarrow c$.
- R2: 反之如果 $a \leftarrow b$，则会形成环 (abc).
- R3：反之如果 $a \leftarrow b$，则 ac 之间的方向必须为 $c \to a$（否则会形成环 (acb)），且 ad 之间的方向也必须为 $d \to a$（否则会形成环 (adb)），而这样的话就会形成新的 V-结构 $c \to a \leftarrow d$.
- R4：反之如果 $a \leftarrow b$，则 ac 之间既不能 $a \to c$（否则会形成环 $(acdb)$），也不能 $a \leftarrow c$（否则会形成 V-结构 $c \to a \leftarrow b$）.

3 PC-Algorithm

PC 算法由 Peter Spirtes and Clark Glymour 与 1991 年提出 [3]，名字取自二人名字首字母。PC 算法的主要贡献是高效实现了 SGS 算法中描述的第 2 步和 IC 算法中所描述的第 1 步。结合 *Causation, Prediction, and Search* 一书中 5.4.2 节的描述 [5]，PC 算法的步骤为：

`Input`：A distribution P or samples obtained from P.

`Output`：A collection of directed acyclic graphs

1. Form the complete undirected graph C on the vertex set V.
2. - $n = 0$
   - `repeat`
     - `repeat`
       * Select an ordered pair of variables X and Y that are adjacent in C such that **`Adjacencies`** $(C, X) \setminus \{Y\}$ has the cardinality greater than or equal to n, and a subset S of **`Adjacencies`** $(C, X) \setminus \{Y\}$ of cardinality n, and if X and Y are d-seperated given S, delete edge $X - Y$ from C and record S in **`Sepset`**(X, Y) and **`Sepset`**(Y, X)
     - `until` all ordered pairs of adjacent variables X and Y such that **`Adjacencies`** $(C, X) \setminus \{Y\}$ has cardinality greater than or equal to n and all subsets S of **`Adjacencies`** $(C, X) \setminus \{Y\}$ of cardinality n have been tested for d-seperation.
     - $n = n + 1$
   - `until` for each ordered pair of adjacent vertices X, Y, **`Adjacencies`** $(C, X) \setminus \{Y\}$ is of cardinality less than n
3. For each triple of vertices (X, Y, Z) such that the pair (X, Y) and the pair (Y, Z) are each adjacent in C but the pair (X, Z) is not adjacent in C, orient $X - Y - Z$ as $X \rightarrow Y \leftarrow Z$ if and only if Y is not in **`Sepset`**(X, Z). Output all graphs consistent with these orientations.

可以看到，这里 Spirtes 的输出依然采用的是所有可能的 DAG，且不包含 IC 算法中第 3 步的内容。而后来在 *Causation, Prediction, and Search* 一书中，Peter Spirtes 为其增加了 IC 算法中的第 3 步的内容，仍未采用 R1-R4 的规则。

4 实现算法

本次实验，我们会采用与 IC 算法相同的输入输出，即输出 PDAG。先采用 PC 算法中的方案构建无向图骨架，后采用 Verma 的四条规则 [7] 标记方向。

`Input:` $\hat{P}$ a stable distribution on a set V of variables.

`Output:` A PDAG C compatible with $\hat{P}$.

1. Form the complete undirected graph C on the vertex set V.
2. - $n = 0$
   - `repeat`
     - `repeat`
       * Select an ordered pair of variables X and Y that are adjacent in C such that **`Adjacencies`** $(C, X) \setminus \{Y\}$ has the cardinality greater than or equal to n, and a subset S of **`Adjacencies`** $(C, X) \setminus \{Y\}$ of cardinality n, and if X and Y are d-seperated given S, delete edge $X - Y$ from C and record S in **`Sepset`**(X, Y) and **`Sepset`**(Y, X)
     - `until` all ordered pairs of adjacent variables X and Y such that **`Adjacencies`** $(C, X) \setminus \{Y\}$ has cardinality greater than or equal to n and all subsets S of **`Adjacencies`** $(C, X) \setminus \{Y\}$ of cardinality n have been tested for d-seperation.
     - $n = n + 1$
   - `until` for each ordered pair of adjacent vertices X, Y, **`Adjacencies`** $(C, X) \setminus \{Y\}$ is of cardinality less than n
3. For each triple of vertices (X, Y, Z) such that the pair (X, Y) and the pair (Y, Z) are each adjacent in C but the pair (X, Z) is not adjacent in C, orient $X - Y - Z$ as $X \to Y \leftarrow Z$ if and only if Y is not in **`Sepset`**(X, Z).
4. Recursively adding arrowheads according to the following three rules:
   - R1: Orient $b - c$ into $b \to c$ whenever there is an arrow $a \to b$ such that a and c are nonadjacent.
   - R2: Orient $a - b$ into $a \to b$ whenever there is chain $a \to c \to b$.
   - R3: Orient $a - b$ into $a \to b$ whenever there are two chains $a - c \to b$ and $a - d - \to b$ such that c and d are nonadjacent.

参考文献

[1] MEEK, C. Causal inference and causal explanation with background knowledge. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (San Francisco, CA, USA, 1995), UAI'95, Morgan Kaufmann Publishers Inc., p. 403–410.

[2] PEARL, J. Causality: Models, Reasoning and Inference, 2nd ed. Cambridge University Press, USA, 2009.

[3] SPIRTES, P., AND GLYMOUR, C. An algorithm for fast recovery of sparse causal graphs. Social science computer review 9, 1 (1991), 62–72.

[4] SPIRTES, P., GLYMOUR, C., AND SCHEINES, R. Causality from probability.

[5] SPIRTES, P., GLYMOUR, C., AND SCHEINES, R. Causation, Prediction, and Search, 2nd Edition, vol. 1 of MIT Press Books. The MIT Press, February 2001.

[6] VERMA, T., AND PEARL, J. Equivalence and synthesis of causal models. In Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence (USA, 1990), UAI '90, Elsevier Science Inc., p. 255–270.

[7] VERMA, T., AND PEARL, J. An algorithm for deciding if a set of observed independencies has a causal explanation. In Proceedings of the Eighth International Conference on Uncertainty in Artificial Intelligence (San Francisco, CA, USA, 1992), UAI'92, Morgan Kaufmann Publishers Inc., p. 323–330.