Conditional independence (d-separation)

21.04.21 09:28

- motivation: what is "blocking"? what happens when we condition on colliders? what is d-separation?

Blocking: (on a chain)

- a path in a DAG can be blocked by conditioning on nodes in the path

$\underbrace{\boxed{A \to G \to B}}_{\text{here: } A \not\perp\!\!\!\perp B}$ ↳ A & B are associated (marginally)

: in this path, if we condition on G (a node in the middle of the path (chain)), we block the path from A to B (path becomes inactive)

- "conditioning on G" is another way to say "controling for G

⇒ blocking G means to make this node's value the same in whole population (i.e. we control for G)

Blocking on a fork:

$A \leftarrow G \rightarrow B$ (→ A associated with B because G affects both A and B)

↳ if we condition on G (i.e. control for G) then the path btw. A and B is blocked.

Colliders:

- opposite to forks
- assume inverted fork path: $A \to G \leftarrow B$
  - ↳ here: $A \perp\!\!\!\perp B$ (because information collides at G)
  - ↳ however: conditioning on G creates/induces an association btw. A and B.

Conditioning on colliders:

example: G: lightbulb is lit up (on, off)
A: state of switch 1 (on, off)
B: state of switch 2 (on, off)
} their values are independently determined by a coin flip

⇒ assume G is only lit up only if both switches A and B are in the on state.

⇒ here A is totally independent of B b/c they are based on two independent coin flips.

→ relationship as DAG: $A \longrightarrow G \longleftarrow B$

↳ here $A \perp\!\!\!\perp B$ (b/c if we know A is on, it doesn't give us any clue whether or not B is on.)

BUT: $A \not\perp\!\!\!\perp B \mid G$ (b/c if we know that G is lit up, then we have to conclude that both switches, A, B, are in the on-state)

↓

also: • if G is off, then assuming that B is in the on-state, we have to conclude that A is in the off-state.
• if G is off, then assuming A is on, we conclude/require B to be off.

When we condition on the collider G, we actually open/create a path between A and B:

$A \to \boxed{G} \leftarrow B$ (new path due to collider conditioning) ⇒ $A \not\perp\!\!\!\perp B \mid G$

grey: conditioning (means that we either fix G or we observe the value it took on)

D-separation (D... dependence):

- a path is d-separated by a set of nodes C if: (*)

• assume we have a path btw. 2 nodes D and F, which are dependent due to possible info flow btw. them.

↳ now we ask: is there a set of nodes which creates independence btw. D and F.

→ if a set of vars C in the path causes D to be independent from F, then we call it a d-separated path (by the set C).

(*) (path) it contains a chain and the middle part is in C. (OR)
↳ e.g. $D \to E \to F$, then $C := \{E\}$

(*) (path) it contains a fork and the middle part is in C. (OR)
↳ e.g. $D \leftarrow E \to F$, then $C := \{E\}$

(*) (path) it contains an inverted fork and the middle part is NOT in C, nor are any descendants of the middle part (collider) in C.
↳ e.g. $D \to E \leftarrow F$, then $C := \{\dots\} \setminus E$

⇒ the idea of d-separation is to make 2 end-nodes in a path independent by blocking/conditioning on C.

d-separation with many paths btw. 2 nodes in the DAG:

→ 2 nodes, A and B, are d-separated by a set of nodes C iff the set C blocks all/every paths from A to B.

⇒ achieve conditional independence btw. A and B through d-separation
↳ $A \perp\!\!\!\perp B \mid C$

↳ recall ignorability assumption: $\underbrace{Y^0, Y^1 \perp\!\!\!\perp A \mid X}$

goal: to identify set of nodes X, that will create conditional independence between A and potential outcomes

task: find set X which d-separates A from $Y^{(i)}$ $\forall i \in A$.

implies d-separation btw. A and the potential outcomes $Y^{(i)}$.