Causality 9/27 10-701

Directed Acyclic Graphs (DAGs)

"conditional independence diagrams"
"Bayesian networks"
"Bayes Nets"

Formally: specifies a set of (known? assumed?) conditional independencies

Traditional view:
graph is a "notational device" for visualizing
cond independencies
inference algorithms

Recap
$$P(X_1, X_2, \ldots X_n)$$
$$= \prod_{i=1}^{n} P(X_i \mid \text{all } X_{<i})$$ ← chain rule of prob
$$= \prod (X_i \mid Pa(X_i))$$

Big parameter savings by taking graph into account.

- Enables tractable inference. *
- provides inductive bias

But why should we believe in this model?
When might we expect cond ind to hold?

What differentiates these graphs?

$$X \rightarrow Y$$

$$X \leftarrow Y$$

Or these

$$X \rightarrow Y \rightarrow Z$$

$$X \leftarrow Y \leftarrow Z$$

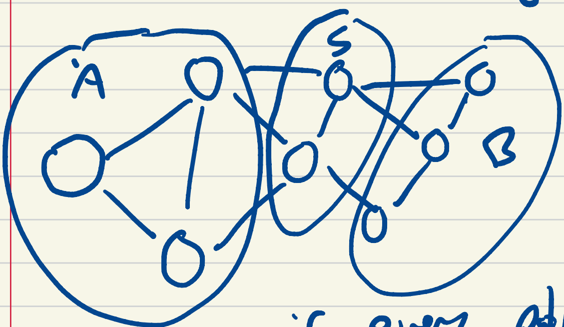
$$X \leftarrow Y \rightarrow Z$$

$$X \rightarrow Y \leftarrow Z$$

Markov Random Fields

DAGs not the only language for expressing cond ind

- Undirected graphical model
- Common in image processing



Set of vars A cond ind of B given S

$A \perp B \mid S$

if every path from A to B goes through S.

What's more expressive: DAG or MRF?

Write fully ind DAG/MRF

Write fully dep DAG/MRF

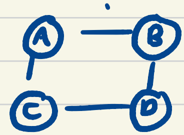
Write equiv

DAG

?

MRF

?



ML - learning from data + assumptions

Stat learning - assumptions on functional forms
hypothesis classes

Causality will introduce new family of
statements + corresponding assumptions

What is the flavor of different kinds of
assumptions? What must I believe about the world?

Arrange vars in a graph:

(1) Cancer
Smoking
tar

possible gene that causes smoking + also predisposes to cancer

(2) Earthquake
Burglary
Alarm

(3) improvise

Causality

Who's heard of it?
What does it mean to you?

(discuss)

Causality concerns interventions

$$\vec{X}, Y \longrightarrow X, T, Y$$

$$Y \mid X \longrightarrow Y \mid do(T=t), X$$

Why is Causality important?

- Perhaps not in Prediction
  (or is it?! dataset shift; how to construct Bayes net)

Often (more than we admit) we want data to tell us what to do

Causality's importance (cont'd)

Healthcare

- personalized medicine
- risk scoring
  (you get a nice prediction → so what?)

Carnema example (be careful)

$$X \rightarrow T \rightarrow Y$$

Asthma treatment death

Other Domains: (i) hiring (ii) recommendation

Primary Areas of Causality

- Inference - Answer questions abt effects of interventions
- Discovery - infer graph from data (as much as possible)

New wrinkle: Identification

History

1739 - Hume's two definitions

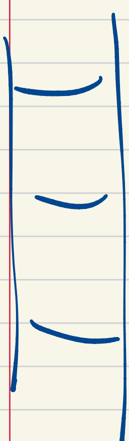
1918 - Sewall Wright's path diagrams

1940s-70s SEMs enter social science

1993 - Formulation as DAGs
Spirtes, Glymour, Scheines

1995 - Pearl identification results
SCM formalized
"do"-calculus

Ladder of Causation

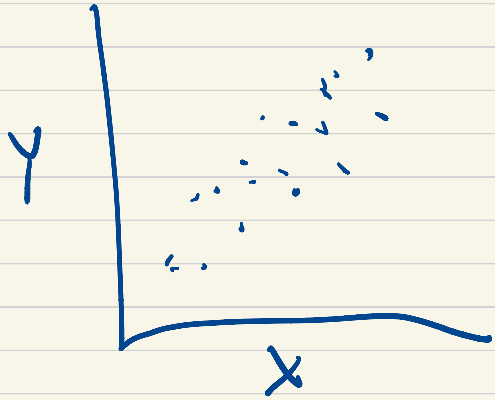


3 Imagining (Counterfactuals)

2 Doing (Intervention)

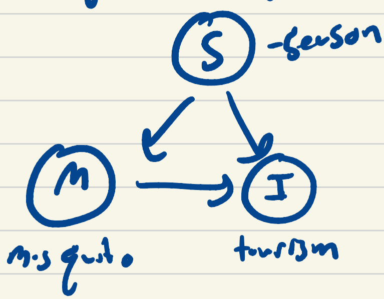
1 Seeing (Association)

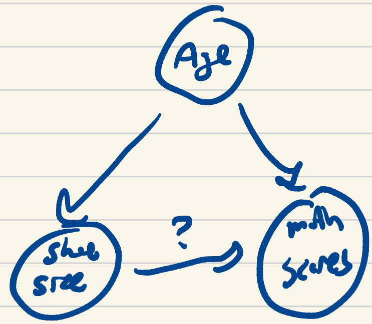
Why might causality $\neq$ Correlation?

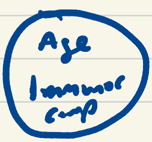


does X cause Y?

Simpson's paradox







Heart disease meds

Heart attack?

Confounding

Central problem of causality

Common causes of both treatment & outcome

How to eliminate Confounding?

Randomized Controlled Trials (RCTs)

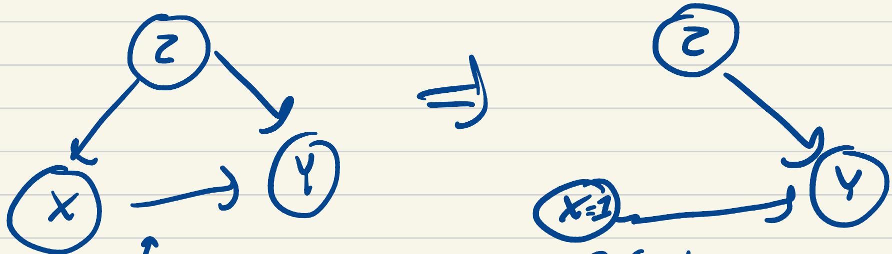
Benefits - in resulting data conditional & interventional coincide

Cons - Not always possible or ethical
smoking, interest rates

Two Notational Worlds
- Graphs (CMU/Pearl)
- Potential outcomes

Technically equivalent, but different attitudes, emphasis, flavor.

Intervention as edge deletion



$P(Y \mid X)$ $\qquad$ $P(Y \mid do(X=1))$

Causal interpretation of graph
direct edges, chains, etc

SCM $\quad$ graph + structural equations/conditional distributions

Causal Quantities of Interest

Average Treatment Effect

$$P(Y \mid do(T=1)) - P(Y \mid do(T=0))$$

Conditional Avg Tr. Eff.

$$P(Y \mid do(T=1), X) - P(Y \mid do(T=0), X)$$

Counterfactuals

Back-door adjustment ✓

$$P(Y \mid do(T=1)) \qquad \sum_X P(Y \mid T=1, X)\, P(X)$$

Front door

Rubin-Neyman Causal Model

$Y_0(x_i)$ — Potential outcome had unit i received treatment 0

$Y_1(x_i)$ — "

* Problem - only one observed

$$Y_i = t_i Y_1(x_i) + (1 - t_i) Y_0(x_i)$$

ATE: $\mathbb{E}[Y_1] - \mathbb{E}[Y_0]$

under RCT $= \mathbb{E}[Y \mid T=1] - \mathbb{E}[Y \mid T=0]$

How to proceed w obs data?

- SUTVA
- Ignorability
- Positivity / common support
- Consistency

$$\mathbb{E}[Y_1] = \mathbb{E}_{x \sim p(x)}\left[\mathbb{E}_{Y_1 \sim p(Y_1|X)}[Y_1 | X]\right]$$

(ignorability) $$= \mathbb{E}[\mathbb{E}[Y_1 | X, T=1]]$$

$$= \mathbb{E}[\mathbb{E}\, Y | X, T=1]]$$

estimable from data

Methods
- Covariate adj
- matching
- propensity score reweighting $\#(T=1|X)$

Outcome regression
"parametric g-formula"

[Explicitly model relationship between outcome and treatment, confounders]

Under ignorability $ATE = E[Y \mid T=1, X] - E[Y \mid T=0, X]$

(1) Fit model $f(x,t) \approx E[Y \mid T=t, x]$

$$\hat{ATE} \quad \sum_{i=1}^{n} f(x_i, 1) - f(x_i, 0)$$