CRG 17/18 Meeting 1:
Intro to Spirtes Glymour & Scheines (2000)

Discussant: Oisín Ryan

March 23, 2018

The Material

Causation, Prediction and Search (2001) by Spirtes, Glymour & Scheines. Second edition

First edition and reprint (1993, 2011) are different from second edition (2001):

- Much longer 1st chapter in 1st edition
- Different presentation of d-seperation in chapter 2 (misleading in first edition)
- New 12th chapter regarding cyclic graphs and feedback systems

For consistency lets stick to the 2nd edition

The Material

Causation, Prediction and Search (2001) by Spirtes, Glymour & Scheines. Second edition

First edition and reprint (1993, 2011) are different from second edition (2001):

- Much longer 1st chapter in 1st edition
- Different presentation of d-seperation in chapter 2 (misleading in first edition)
- New 12th chapter regarding cyclic graphs and feedback systems

For consistency lets stick to the 2nd edition

- Apart from some brief parts of this presentation!

The Plan

- Week 1: Introduction and Chapter 1
- Week 2: Chapter 2
- Week 3: Chapter 3 (3.0 - 3.6)

Preface

Fundamental arguments:

- Return to a vision of statistics with the goal of making causal inferences or predicting effects of manipulations
- Arguments against inferring causes from statistics outside of experimental trials are unsound
- Experimental and observational design are subject to uniform principles

Preface

The theory:

- Two axioms relating casual structures and probability distributions
- Leads to asymptotically reliable search procedure (PC algorithm)
- Shows that current methods (e.g. regression model selection) are "radically suboptimal"
- Clarifies diverse topics: Simpson's paradox, experiment vs observation, errors in regression, retrospective vs prospective, variable selection

Notation

Variables:	capitalized, and in italics, e.g., X
Values of variables:	lower case, and in italics, e.g., $X = x$
Sets:	capitalized, and in boldface, e.g., $\mathbf{V}$
Values of sets of variables:	lower case, and in boldface, e.g., $\mathbf{V} = \mathbf{v}$
Members of $\mathbf{X}$ that are not members of $\mathbf{Y}$:	$\mathbf{X}\backslash\mathbf{Y}$
Error variables:	ε, δ, e
Independence of $\mathbf{X}$ and $\mathbf{Y}$:	$\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$
Independence of $\mathbf{X}$ and $\mathbf{Y}$ conditional on $\mathbf{Z}$:	$\mathbf{X} \perp\!\!\!\perp \mathbf{Y}\mid\mathbf{Z}$
$\mathbf{X} \cup \mathbf{Y}$:	$\mathbf{XY}$
Covariance of X and Y:	$\mathrm{COV}(X,Y)$ or γ_{XY}
Correlation of X and Y :	ρ_{XY}
Sample correlation of X and Y :	r_{XY}
Partial Correlation of X and Y, controlling for all members of set $\mathbf{Z}$:	$\rho_{XY.\mathbf{Z}}$

Notation

Variables:	capitalized, and in italics, e.g., X
Values of variables:	lower case, and in italics, e.g., $X = x$
Sets:	capitalized, and in boldface, e.g., $\mathbf{V}$
Values of sets of variables:	lower case, and in boldface, e.g., $\mathbf{V} = \mathbf{v}$
Members of $\mathbf{X}$ that are not members of $\mathbf{Y}$:	$\mathbf{X} \backslash \mathbf{Y}$
Error variables:	ε, δ, e
Independence of $\mathbf{X}$ and $\mathbf{Y}$:	$\mathbf{X} \perp\!\!\!\perp \mathbf{Y}$
Independence of $\mathbf{X}$ and $\mathbf{Y}$ conditional on $\mathbf{Z}$:	$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \mid \mathbf{Z}$
$\mathbf{X} \cup \mathbf{Y}$:	$\mathbf{XY}$
Covariance of X and Y:	$\mathrm{COV}(X,Y)$ or γ_{XY}
Correlation of X and Y :	ρ_{XY}
Sample correlation of X and Y :	r_{XY}
Partial Correlation of X and Y, controlling for all members of set $\mathbf{Z}$:	$\rho_{XY.\mathbf{Z}}$

Notation

Notation, proofs, results given for discrete variables

Generalises to continuous variables

- Probability distribution $\rightarrow$ Density function
- Summations $\rightarrow$ integrals

$\sum^{\rightarrow}$ - sum over values of the random variables

- For dichotomous X:
  $\sum_X^{\rightarrow} P(X|Y=0) = P(X=0|Y=0) + P(X=1|Y=0)$

Intro

Many statisticians avoid explicit discussion of causality because:

- Causal claims have complexity and variety
- Claims about what did not happen, or *what would have happened* if some circumstance was changed

However most research is concerned with causal relationships

- Predict the effects of strategies/treatments

Intro

Traditionally missing a rigourous theory of causal inference from non-experimental observations

Conditioning is not intervening:

- In many causal systems the probability of an event Y given an intervention to bring about an event X is different from the conditional probability of Y on X.

Intro

Three problems:

1. Clarifying the notion of a causal system with enough precision
2. Understanding possibilities and limitations for discovering such causal structures from different types of data
3. Characterizing probabilities predicted by a causal hypothesis given an intervention on variables

Intro

Approach

1. Using graphical formalism of Speed, Pearl and others
2. Discovery with algorithms developed from the mathematics of this graphical representation
3. Outline the theory of manipulation and its assumptions
   - Put these assumptions in a graphical framework

Manipulation and graphical approaches to causal inference

Counterfactual approaches - Rubin, Holland and others

- Missing data problem
- Causal hypotheses postulate a family of random variables, some of which **never** have their values observed
- What *would* have happened if everyone had received the treatment vs everyone received the control
- Specify models for treatment assignment and models for counterfactual outcomes
- Definition of causal effects in terms of hypothetical experiments

Graphical approaches - Spirtes, Glymour & Scheines, Pearl, Lauritzen

- Factorisation of joint densities with conditionals
- Connection between Directed Graphs, Markov conditions and probability densities
- Less focus on definition of causal effects - probability is also vague and axiomatic
- Both Spirtes et al and Pearl claim Rubin framework as special cases

Intro

Directed Graphical Models reflect two fundamental causal notions

1. Absence of causal relation $\rightarrow$ independence in probability
2. Probability is associated with control
   - Variation in X causes variation in Y $\rightarrow$ Y can be changed by altering X

Aim to characterize

- when alternative causal theories are indistinguishable by data
- which features are shared by all indistinguishable models