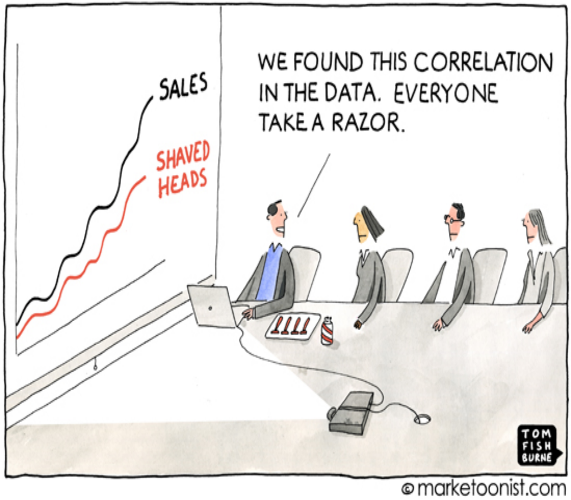
A Gentle Introduction to Structural Causal Models

Daniel Saggau

5/11/2021

Motivation


SALES
SHAVED HEADS
WE FOUND THIS CORRELATION IN THE DATA. EVERYONE TAKE A RAZOR.
TOM FISH BURNE
© marketoonist.com

Introduction (1)

- Machine learning has provided many insights into different problems
- One issue is the consideration of ‘What are we actually predicting?’
- Mainstream tools are build on association-based learning
- Associations are not enough for high stake settings
- In disciplines like psychology or economics people are less interested in associational learning
- We want causation and not correlation

Introduction (2)

Causal assumptions differentiate causal models from association learning methods.

Association-based Concepts	Causal Concepts
Correlation	Randomization
Regression	Confounding
Conditional Independence	Disturbance
Likelihood	Error Terms
Odds Ratio	Structural Coefficients
Propensity Score	Spurious Correlation

Table of Contents

Assumptions

- Consists of system of equations
- Assignment equation ‘$:=$’ rather than regular equation ‘$=$’
- is a nonparametric SEM
- has functional form rather than using probabilities
- entails features from the PO framework and graphical representation
- Exogenous factors are part of the model specification

Assumptions

Error terms

Regression: Omittable outside factor SCM/SEM; Latent influential factor that is pivotal for the model specification but not observable

Consists of graph and assignments: Baseline:

$$C := N_c$$

$$E := f_E(C, N_E)$$

source: Peters, Janzing, and Schölkopf (2017)

SCM Applications:

- Flexible simulations for higher order problems (intervention, counterfactual)
- Graphical visualization via directed acyclic graph

Comparative Causal Tools

Historical Development

- Path Analysis -> SEM -> SCM

Fundamental Differences (1)

- conflict whether to use graphs or not
- A SEM is a parametric specification used in applied sciences (parameters contested)
- A Bayesian causal network is another popular causal model using conditional probabilities and NO functions
- Differences in performance between BCN and SCM# Performance Evaluation

Implications

Method	CBN	SCM
Prediction	• Unstable • Volatile to parameter changes • Re-Estimate entire model	• Stable • More Natural Specification • Only estimate Δ CM
Intervention	• Costly for Non-Markovian Models • Unstable(Nature CP) • Only generic estimates(Δ CP)	• Pot. Cyclic Representation • Stable(Nature Eq.) • Context specific(Invariance of Eq.)
Counterfactuals	• **Impossible** • no information on latent factors(ϵ)	• Possible • Inclusion of latent factors

Pearls Causal Hierachy

Table 3: Pearls Hierachy of Causation (2009)

Method	Action	Example	Usage
Association $P(a\|b)$	Co-occurrence	What happened...	(Un-)Supervised ML, BN, Reg.
Intervention $P(a\|do(b), c)$	Do-manipulation	What happens if ...	CBN,MDP,RL
Counterfactual $P(a_b\|a`, b`)$	Hypotheticals	What would have happened if...	SCM ,PO

Graphical Illustration

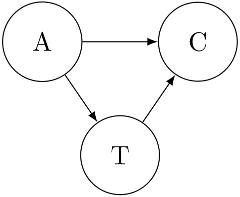


Figure 1: Probabilistic Model

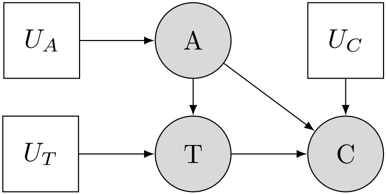


Figure 2: Structural Causal Model

Causality and Time

- Time in Physical Sciences: Mechanical and exact
- Time in Social Sciences: Often Vague
- Regular Time Specification is also more vague
- To accomodate that issue, research on differential equation based SCMs started

References

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. 2017. *Elements of Causal Inference: Foundations and Learning Algorithms*. Adaptive Computation and Machine Learning Series. Cambridge, Massachuestts: The MIT Press.