# Structural Equation Models: Where do they come from?

Paul Rubenstein

University of Cambridge

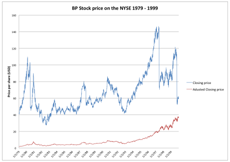
Max-Planck Institute for Intelligent Systems, Tübingen

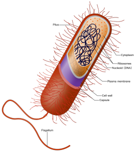
November 22, 2016

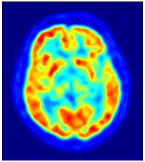
# Outline

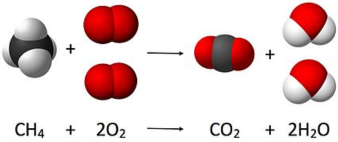
# Motivation

Real world: complex and time evolving



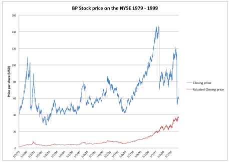



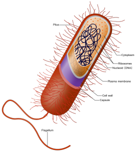



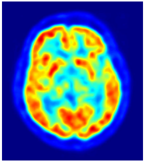


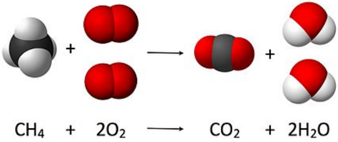
# Motivation

Real world: complex and time evolving








Imperfect measurements

- Destructive
- Indirect
- Slow timescale

# Motivation

Machine Learning: write down 'plausible' model

$$\mathbf{X} \overset{i.i.d.}{\sim} \mathbb{P}_X^{\theta}$$

Learn/estimate parameters $\theta$

# Motivation

Machine Learning: write down 'plausible' model

$$\mathbf{x} \overset{i.i.d.}{\sim} \mathbb{P}_X^\theta$$

Learn/estimate parameters $\theta$

Why learn 'true' generative process?

- Interpretability
- Control

# Motivation

Machine Learning: write down 'plausible' model

$$\mathbf{X} \overset{i.i.d.}{\sim} \mathbb{P}_X^\theta$$

Learn/estimate parameters $\theta$

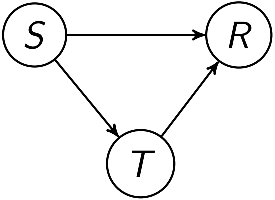
Why learn 'true' generative process?

- Interpretability
- Control

Structural Equation Models

# Structural Equation Models (SEMs)

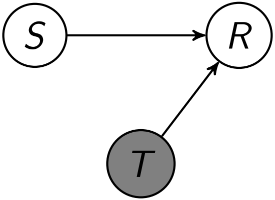
Doctor sees patient with **S**ymptoms, gives **T**reatment and patient **R**ecovers or not.




$\mathbb{P}_{STR}$ determined by

$$\text{Probabilistic Modelling} \left\{ \begin{array}{lcl} \mathbb{P}_S & \longleftrightarrow & S = f_S(E_S) \\ \mathbb{P}_{T|S} & \longleftrightarrow & T = f_T(S, E_T) \\ \mathbb{P}_{R|T,S} & \longleftrightarrow & R = f_R(S, T, E_R) \\ & & \mathbb{P}_{E_S E_T E_R} \end{array} \right\} \text{SEM}$$

# Structural Equation Models (SEMs)

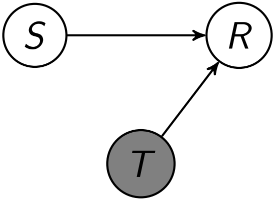



Model | do($T = t$)

$$\text{Probabilistic Modelling} \left\{ \begin{array}{lcl} \mathbb{P}_S & \longleftrightarrow & S = f_S(E_S) \\ \delta(T - t) & \longleftrightarrow & T = t \\ \mathbb{P}_{R|T,S} & \longleftrightarrow & R = f_R(S, T, E_R) \\ & & \mathbb{P}_{E_S E_T E_R} \end{array} \right\} \text{SEM}$$

Perfect interventions override 'causal mechanism' $f$.

# Structural Equation Models (SEMs)




Model | do($T = t$)

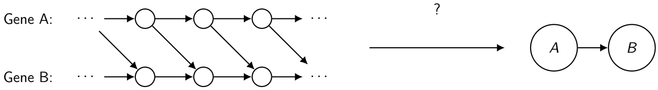
$$\text{Probabilistic Modelling} \left\{ \begin{array}{lcl} \mathbb{P}_S & \longleftrightarrow & S = f_S(E_S) \\ \delta(T - t) & \longleftrightarrow & T = t \\ \mathbb{P}_{R|T,S} & \longleftrightarrow & R = f_R(S, T, E_R) \\ & & \mathbb{P}_{E_S E_T E_R} \end{array} \right\} \text{SEM}$$

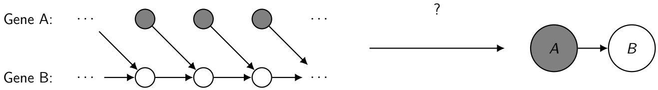
Perfect interventions override ‘causal mechanism’ $f$.

For fixed $\theta$, SEM implies *family* of $\mathbb{P}_{STR}$ indexed by intervention.

# The Problem

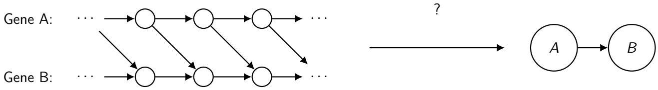
Eg. Cell with gene knockouts, partially observable.

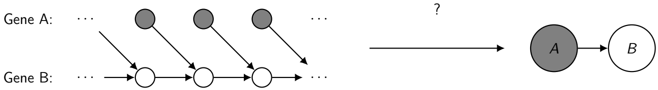


# The Problem

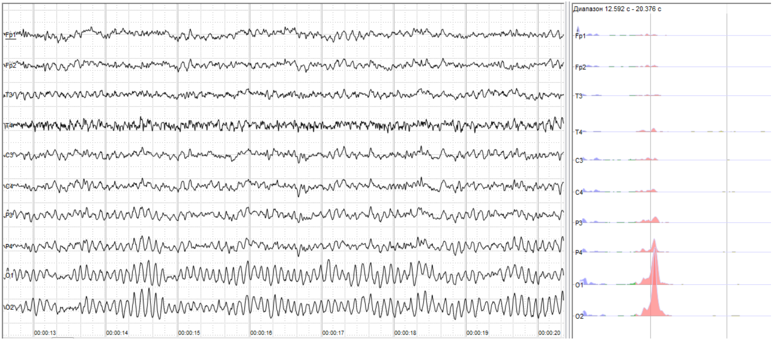
Eg. Cell with gene knockouts, partially observable.

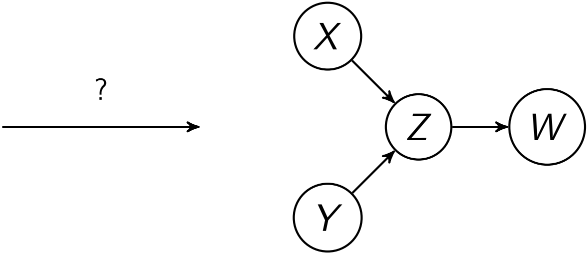



How do we interpret SEMs when variables are 'abstracted' from true generative structure?

# The Problem


Fp1
Fp2
T3
T4
C3
C4
P3
P4
O1
O2
00:00:13
00:00:14
00:00:15
00:00:16
00:00:17
00:00:18
00:00:19
00:00:20
Диапазон 12.592 с - 20.376 с

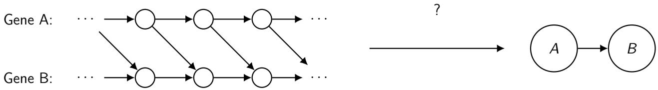

?
X
Z
W
Y

# The Problem

Question:

How do the 'Transformed' SEMs relate to the originals?

How do we make sense of interventions in the Transformed SEMs?



# Prior work

Similar problem considered in Mooij, Janzing and Schölkopf (2013):

Can equilibria of ODEs be described by an SEM?

Consider 'constant' interventions.

# Prior work

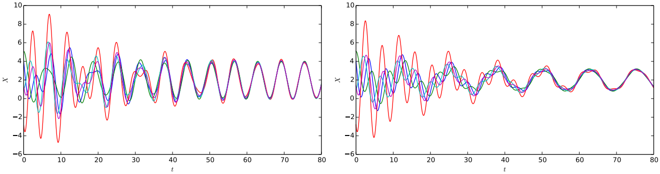
Similar problem considered in Mooij, Janzing and Schölkopf (2013):

Can equilibria of ODEs be described by an SEM?

Consider 'constant' interventions.

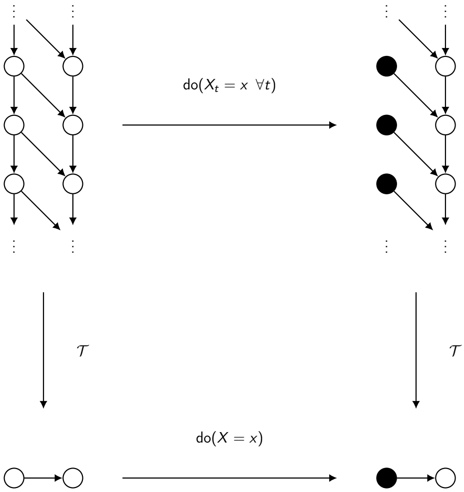
Extended further to include non-constant interventions to ask:

Can asymptotic behaviour of ODEs be described by an SEM?



# The Framework

In brief:



# The Framework

Suppose that $\mathcal{M}_X = \{\mathcal{S}_X, \mathcal{I}, \mathbb{P}_{E_X}\}$ is an SEM

- $\mathcal{S}_X$ = Structural equations on $X \in \mathcal{X}$
- $\mathcal{I}$ = Restricted set of interventions
- $\mathbb{P}_E$ = Distribution over exogenous variables

# The Framework

Suppose that $\mathcal{M}_X = \{\mathcal{S}_X, \mathcal{I}, \mathbb{P}_{E_X}\}$ is an SEM

- $\mathcal{S}_X$ = Structural equations on $X \in \mathcal{X}$
- $\mathcal{I}$ = Restricted set of interventions
- $\mathbb{P}_E$ = Distribution over exogenous variables

$\mathcal{M}_X$ defines observational distribution over $\mathcal{X}$:

$$\mathbb{P}_X$$

For each $i \in \mathcal{I}$, $\mathcal{M}_X$ define interventional distribution over $\mathcal{X}$:

$$\mathbb{P}_X^{\mathrm{do}(i)}$$

Thus $\mathcal{M}_X \implies \mathcal{P}_X = \{\mathbb{P}_X^{\mathrm{do}(i)} \; : \; i \in \mathcal{I}\}$

# The Framework

$$\mathcal{M}_X \implies \mathcal{P}_X = \{\mathbb{P}_X^{\mathrm{do}(i)} \; : \; i \in \mathcal{I}\}$$

# The Framework

$$\mathcal{M}_X \implies \mathcal{P}_X = \{\mathbb{P}_X^{\mathrm{do}(i)} \; : \; i \in \mathcal{I}\}$$

Let $\tau : \mathcal{X} \longrightarrow \mathcal{Y}$

- $\mathcal{Y}$ is 'simpler' or 'transformed' space
- $\tau$ could represent e.g. measurement or transformation to data

# The Framework

$$\mathcal{M}_X \implies \mathcal{P}_X = \{\mathbb{P}_X^{\mathrm{do}(i)} \;:\; i \in \mathcal{I}\}$$

Let $\tau : \mathcal{X} \longrightarrow \mathcal{Y}$

- $\mathcal{Y}$ is 'simpler' or 'transformed' space
- $\tau$ could represent e.g. measurement or transformation to data

$X$ a random variable $\implies$ $\tau(X)$ a random variable.

Define $\mathbb{P}_Y := \tau(\mathbb{P}_X)$ via push-through measure.

Similarly, $\mathbb{P}_Y^i := \tau\left(\mathbb{P}_X^{\mathrm{do}(i)}\right)$

$$\mathcal{P}_Y := \{\mathbb{P}_Y^i \;:\; i \in \mathcal{I}\}$$

# The Framework

$\mathcal{M}_X \implies \mathcal{P}_X = \{\mathbb{P}_X^{\mathrm{do}(i)} \ : \ i \in \mathcal{I}\}$

Let $\tau : \mathcal{X} \longrightarrow \mathcal{Y}$

- $\mathcal{Y}$ is 'simpler' or 'transformed' space
- $\tau$ could represent e.g. measurement or transformation to data

$X$ a random variable $\implies$ $\tau(X)$ a random variable.

Define $\mathbb{P}_Y := \tau(\mathbb{P}_X)$ via push-through measure.

Similarly, $\mathbb{P}_Y^i := \tau\left(\mathbb{P}_X^{\mathrm{do}(i)}\right)$

$$\mathcal{P}_Y := \{\mathbb{P}_Y^i \ : \ i \in \mathcal{I}\}$$

Does there exist $\mathcal{M}_Y = \{\mathcal{S}_Y, \mathcal{J}, \mathbb{P}_{E_Y}\}$ such that $\mathcal{M}_Y \implies \mathcal{P}_Y$?

# The Framework

If YES:

$$\forall i \in \mathcal{I} \; \exists j \in \mathcal{J} \text{ such that } \mathbb{P}_Y^{\mathrm{do}(j)} = \mathbb{P}_Y^{i}$$

# The Framework

If YES:

$\forall i \in \mathcal{I} \; \exists j \in \mathcal{J}$ such that $\mathbb{P}_Y^{\mathrm{do}(j)} = \mathbb{P}_Y^i$

Thus can extend $\tau : \mathcal{I} \longrightarrow \mathcal{J}$ via $\tau(i) = j \iff \mathbb{P}_Y^{\mathrm{do}(j)} = \mathbb{P}_Y^i$

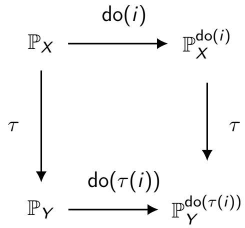
# The Framework

If YES:

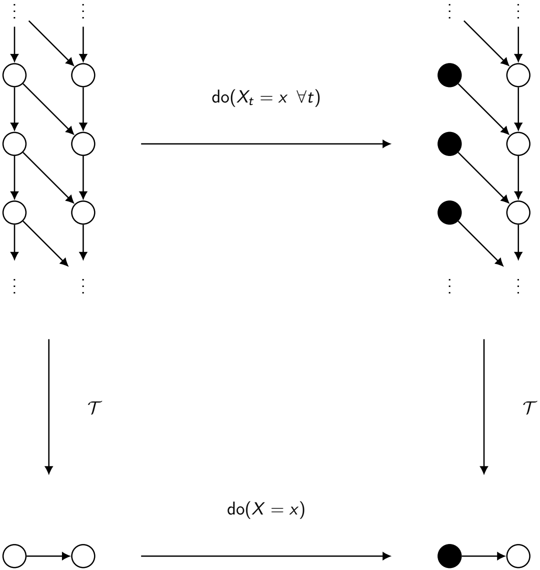
$\forall i \in \mathcal{I} \; \exists j \in \mathcal{J}$ such that $\mathbb{P}_Y^{\text{do}(j)} = \mathbb{P}_Y^i$

Thus can extend $\tau : \mathcal{I} \longrightarrow \mathcal{J}$ via $\tau(i) = j \iff \mathbb{P}_Y^{\text{do}(j)} = \mathbb{P}_Y^i$

Thus $\tau(\mathbb{P}_X^{\text{do}(i)}) = \mathbb{P}_Y^{\text{do}(\tau(i))}$, i.e. the diagram commutes:



# Example



# Some questions

Given $\tau$, what conditions required on $X$?

- i.e. measurement device fixed - when can we represent causal structure?

# Some questions

Given $\tau$, what conditions required on $X$?

- i.e. measurement device fixed - when can we represent causal structure?

Given $X$, what $\tau$ should we apply?

- i.e. how can we massage data to identify causal structure?

# Some questions

Given $\tau$, what conditions required on $X$?

- i.e. measurement device fixed - when can we represent causal structure?

Given $X$, what $\tau$ should we apply?

- i.e. how can we massage data to identify causal structure?

Given $Y$, what $X$ can we abstractly represent?

- i.e. given restricted model class, when can we faithfully represent more complex X?

# Some questions

Given $\tau$, what conditions required on $X$?

- i.e. measurement device fixed - when can we represent causal structure?

Given $X$, what $\tau$ should we apply?

- i.e. how can we massage data to identify causal structure?

Given $Y$, what $X$ can we abstractly represent?

- i.e. given restricted model class, when can we faithfully represent more complex X?

Approximate abstraction by relaxing $\tau(\mathbb{P}_X^{\text{do}(i)}) = \mathbb{P}_Y^{\text{do}(\tau(i))}$

# Summary

Causal structure may exist on levels we cannot observe directly.

Questions:

- Under what conditions can we still represent such structure?
- How can we assign meaning to such representations?

Introduced a framework for analysing transformations

We can now ask:

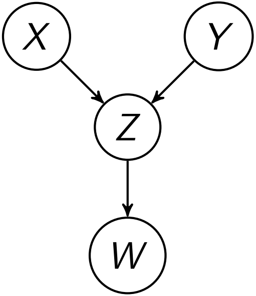
- What conditions required on $X$, $\tau$ and $Y$ for abstraction?

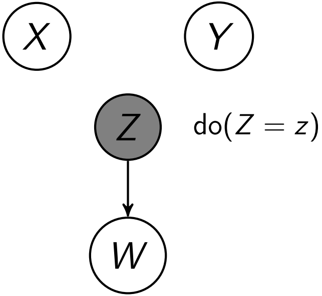
# Fin

Thanks!

# Fin

# Structural Equation Models (SEMs)

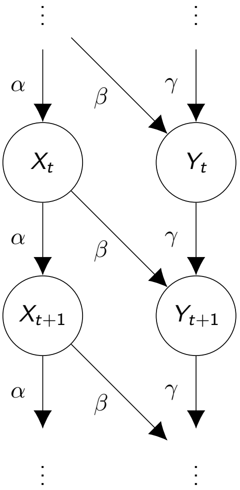



$$
\begin{aligned}
X &= f_X(E_X) \\
Y &= f_Y(E_Y) \\
Z &= f_Z(X, Y, E_Z) \\
W &= f_W(Z, E_W)
\end{aligned}
$$

$$
\begin{aligned}
X &= f_X(E_X) \\
Y &= f_Y(E_Y) \\
Z &= z \\
W &= f_W(Z, E_W)
\end{aligned}
$$

# Example: Linear Gaussian Dynamics



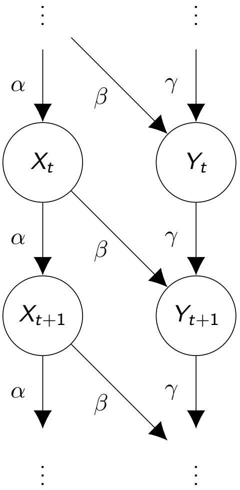

$$X_{t+1} = \alpha X_t + \mathcal{N}(0, \sigma_X^2)$$
$$Y_{t+1} = \beta X_t + \gamma Y_t + \mathcal{N}(0, \sigma_Y^2)$$

Suppose observations come from the stationary distribution

$$\implies \begin{cases} \mathbb{P}_{XY} \\ \mathbb{P}_{Y|\mathrm{do}(X=x)} \\ \mathbb{P}_{X|\mathrm{do}(Y=y)} \end{cases}$$

Problem: unless $\alpha = 1, \sigma_X^2 = 0$, cannot express distributions using unconfounded SEM.

# Example: Linear Gaussian Dynamics



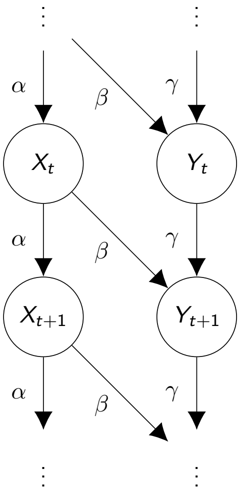

Define 'blurry measurements':

$$X^{(T)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t$$

$$Y^{(T)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y_t$$

As $T \to \infty$, $(X^{(T)}, Y^{(T)}) \xrightarrow{d} (X, Y)$ for which unconfounded SEM can express observational and interventional distributions.

# Example: Linear Gaussian Dynamics




Introduce approximation error:

$$\epsilon = \min_{SEM} \sup_{\zeta \in \{\text{interventions}\}} KL[\mathbb{P}_{\text{do}(\zeta)} | \hat{\mathbb{P}}_{\text{do}(\zeta)}]$$

where $\mathbb{P}_{\text{do}(\zeta)}$ is from true model, $\hat{\mathbb{P}}_{\text{do}(\zeta)}$ is from SEM approximation.

Then $\epsilon < \infty$, and $\epsilon \to 0$ as $\alpha \to 1$, $Var(X) \to c$