Causal Inference

Internship proposal

Internship proposal

In today's technology-focused economy, data are leveraged to model, explain, predict and inform decision-making. However, most models draw upon statistical co-occurences, which may lead to misleading conclusions.

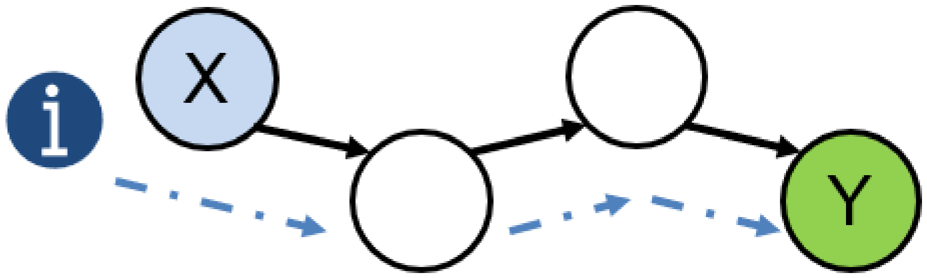
The present internship offers the experience necessary to get a foot in the fascinating domain of causal inference, which is the foundation for developing complex AI architectures based on causal links and resembling human reasoning.

Causality

- The very concept of **causality is questionable**;
- Causality is *epistemologically* related to **applying an intervention** and observing an effect;
- Without **experimental control**, we can only observe *relations* among events, but no causal link;
- However, several methodologies allow to estimate causal connections from:
  - A set of assumptions on data distributions and connection types
  - Prior knowledge
  - Estimated (in)dependencies
  - Estimated conditional independencies

Causal discovery

- The hard definition of causality relies on experimental control
  - Therefore, such a definition is tough to apply to economy-related domains;
  - most dynamics take place out of human control in economy;
- A softer definition of causality allows to individuate **information flow** from X → O → Y
  - e.g.: how macroeconomic factors (X) relate to each other and modulate default rates (Y)



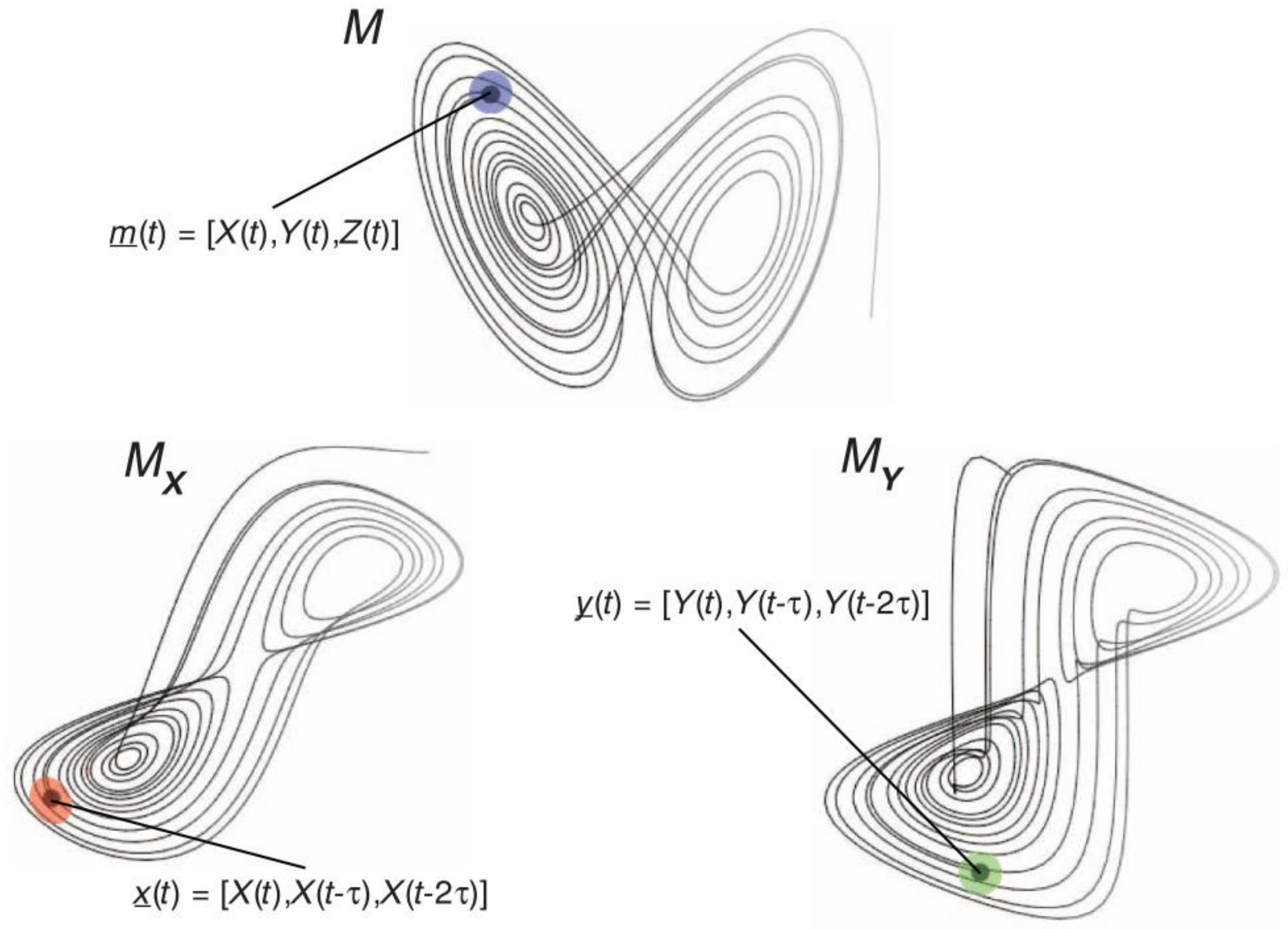
$\therefore X \not\perp\!\!\!\perp Y$

Causal inference

assumption	Causal inference can only occur under quite constricting assumptions. Depending on experimental design and data availability an appropriate set of assumptions must be chosen (if existing).
conforming data	Data must be conformed to model assumptions, in relation to prior knowledge (if any). For example, assumptions may require the relation among variables be linear and the residual of such a relation be non-Gaussian. Therefore, variables not conforming to assumptions must undergo appropriate transformations.
causal discovery	If the causal dynamics are not already known, several algorithms can be leveraged to uncover causal relations. Each algorithm makes several assumptions and expect data to correspond to certain schemas (e.g. time-series, cross-sectional, etc.).
probabilistic graph	The causal graph must be converted into a probabilistic graph which specifies distributional shapes, parameters and priors for both each element and relation in the graph, along with dependence and conditional independencies.
estimation	Parameters in the graph undergo estimation through MCMC or VI.
inference	Depending on the experimental question, a target posterior distribution can be obtained, or the estimated distributions of the elements of the graph is used to compute causal estimates (e.g. ATE).

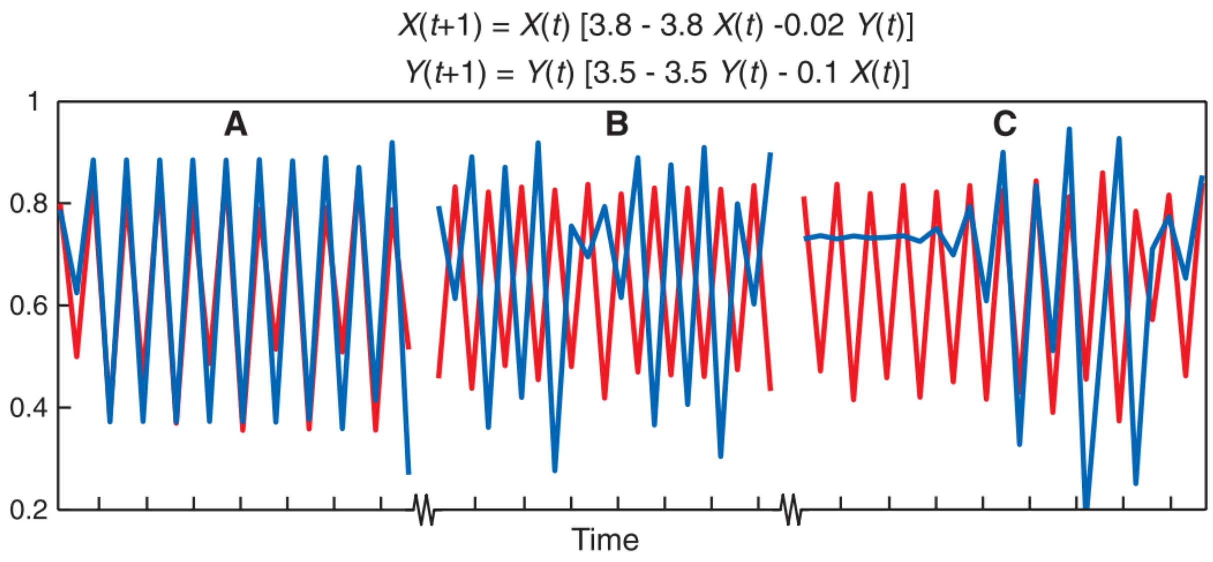
Takens' theorem

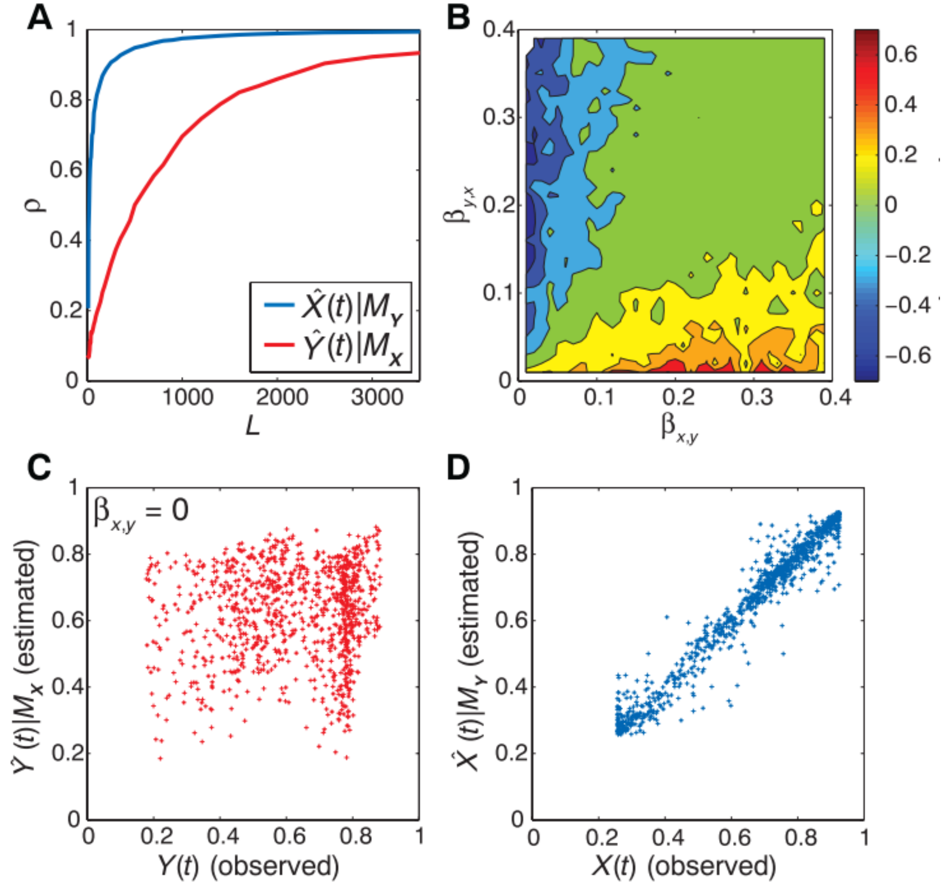
- In dynamic systems with behaviors that are at least somewhat deterministic, information about past states is carried forward through time.
- If x influences y, then the historical values of x can be recovered from variable y alone.
- A time delay embedding is constructed from the time series of y, and the ability to estimate the values of x from this embedding quantifies how much information about x has been encoded into y.



*Electricity, gas, steam and air conditioning supply

Convergent cross mapping





Mirage correlations can emerge from dynamically coupled systems, alternating periods of positive, negative and zero correlation.

However the causal association of X on Y can still be recovered along the whole timeseries.

Complex systems

In principle CCM can recover the dynamics of the system depicted in the figure:

- Here components A, B, C show bidirectional causality;
- in turn A, B and C generate a factor which influences D and E, while D and E don't interact one another.