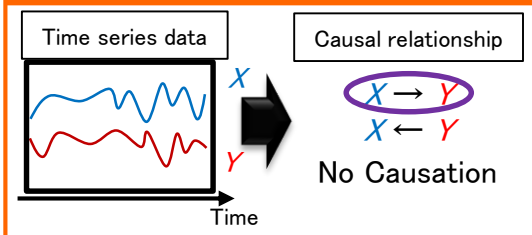
05 Which is cause? Which is effect? Learn from data!

Causal inference in time series via supervised learning

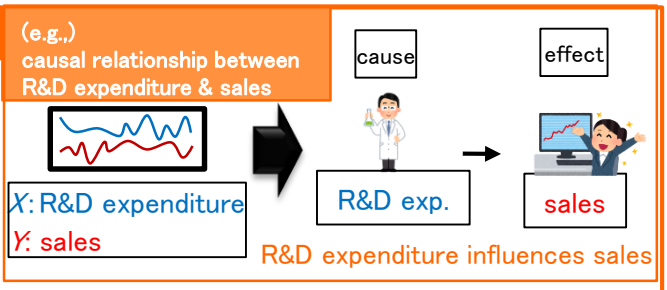
Abstract

Our goal is to automatically discover "causal relationships" from time series data, i.e., a sequence of data measured over time. Discovering causal relationships has key applications in various fields: e.g., finding that "R&D expenditure influences sales" is useful for decision making in companies; discovering gene regulatory relationships provides a key insight for drug discovery researches.
To infer causal relationships, existing methods require us to select an appropriate mathematical expression (i.e., auto-regressive model) for each time series data, which is difficult without expertise in data analysis. For this problem, we build a novel approach that trains a machine learning model by using various data. Our method does not require a deep understanding of data analysis and therefore will help us to effectively make an important decision making in several situations.

Problem setup: causal inference in time series



(e.g.,)
causal relationship between R&D expenditure & sales



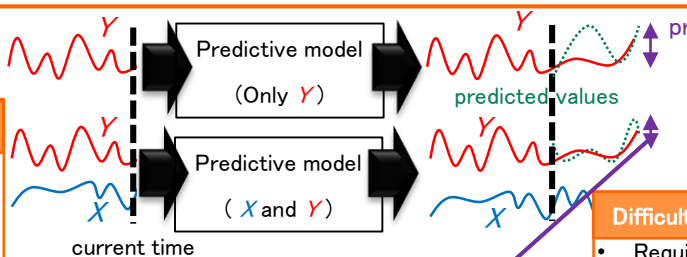
Existing methods

What is causality?

If past values of X are helpful to predict future values of Y, then

$X \rightarrow Y$

(Granger causality)



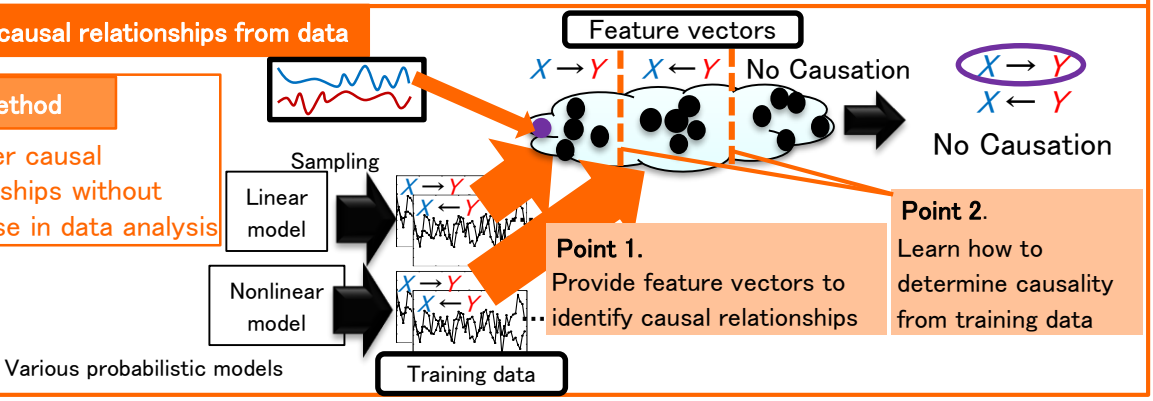
Difficulty

- Require to appropriately select predictive models
- Difficult without expertise in data analysis

Learn causal relationships from data

Our method

can infer causal relationships without expertise in data analysis



References

[1] Y. Chikahara, A. Fujino, "Causal Inference in Time Series via Supervised Learning," in *Proc. 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.

Contact

Yoichi Chikahara Email: cs-liaison-ml at hco.ntt.co.jp
Learning and Intelligent Systems Research Group, Innovative Communication Laboratory

Innovative R&D by NTT
Open House 2019

Copyright © 2019 NTT. All Rights Reserved.