AI CENTER FEE CTU

CoDiet

ECML PKDD 2024

Applications of System Identification

Xiaoyu He, Petr Ryšavý, Mengjia Niu, Quan Zhou and Jakub Mareček

Artificial Intelligence Center
Faculty of Electrical Engineering
Czech Technical University in Prague

Sep 09, 2024

Problem Preliminaries

A **Linear Dynamic System L** is defined as a quadruple **(G, F, Σ_H , Σ_O)**. A single realization of the LDS or a trajectory of length T can be denoted by **X** = $\{x_1, x_2, \ldots, x_T\}$ $\in \mathbb{R}^{m \times T}$.

Based on initial conditions ϕ_0, and realization of noises υ_t and ω_t, LDS is defined as

$$\begin{aligned} \phi_t &= G\phi_{t-1} + \omega_t, \qquad & \omega_t &\sim N(0, \Sigma_H), \\ x_t &= F'\phi_t + \upsilon_t, \qquad & \upsilon_t &\sim N(0, \Sigma_O), \qquad t \in \{1, 2, \ldots, T\}, \end{aligned}$$

where

- G and F are system matrices of dimension n × n and n × m, respectively.
- $\Sigma_H \in \mathbb{R}^{n \times n}$ *and* $\Sigma_O \in \mathbb{R}^{m \times m}$ are covariance matrices.
- $\phi_t \in \mathbb{R}^n$ is the vector autoregressive processes with hidden components.
- The transpose of F is denoted as F'.
- Vector $x_t \in \mathbb{R}^m$ serves as an observed output of the system.

NCPOP

- To find the global optima of the objective function subject to the feasible constraints arising from previous LDS(Q, Zhou and J, Mareček).

$$\min_{f_t,\phi_t,G,F,\omega_t,\upsilon_t} \sum_{t\in\{1,2,\ldots,T\}} \|X_t - f_t\|_2^2 + \|\omega_t\|_2^2 + \|\upsilon_t\|_2^2,$$

for a L_2-norm $\|\cdot\|_2$.

- The optimal objective values can be bounded in non-convex mixed-integer nonlinear programs (MIP). Then they extended the research to a Non-Commutative Polynomial Optimization (NCPOP), which is an operator-valued optimization problem. Given the input data p(x) and $q_i(x)$, the standard form is

$$\begin{aligned} \underset{(\mathcal{H},X,\psi)}{\text{minimize}} \quad & \langle \psi, p(X)\psi \rangle \\ \text{subject to} \quad & q_i(X) \succcurlyeq 0,\ i = 1,\ldots,m, \\ & \langle \psi, \psi \rangle = 1, \end{aligned}$$

where $q_i(x)$ are positive semidefinite.

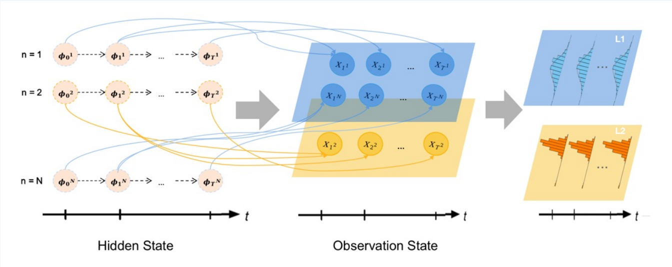
AI Center FEE CTU

Q. Zhou and J. Mareček, "Learning of linear dynamical systems as a non-commutative polynomial optimization problem," *IEEE Transactions on Automatic Control,* 2023.

P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, pp. 1-131, 2013.

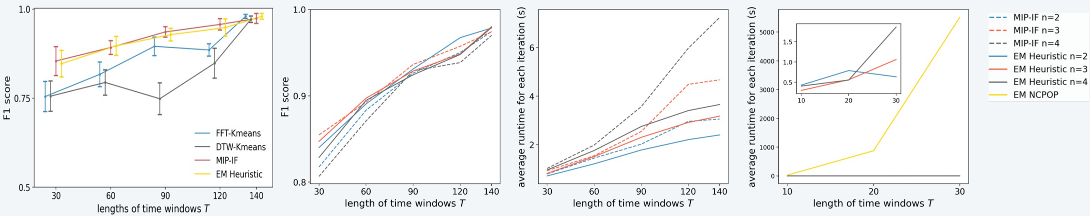
Joint Partition

Clustering of time series is a well-studied problem. Here we consider a variant, where given a set of trajectories and a number of parts, we jointly partition the set of trajectories and learn linear dynamical system (LDS) models for each part, so as to minimize the maximum error across all the models.



EM Heuristic

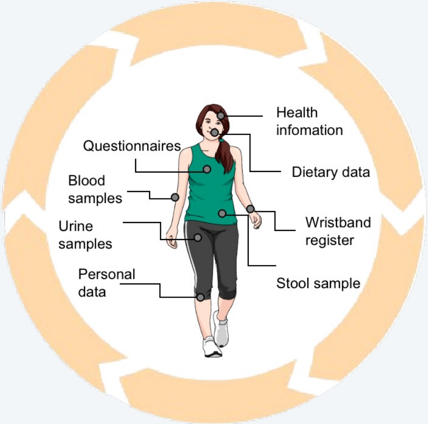
- As a complement to the NCPOP formulation, we provide an efficient Expectation-Maximization (EM) procedure for clustering time series.
- Our approach based on system identification consistently outperforms conventional methods, particularly as the length of the time window increases.



AI Center FEE CTU

- **Spain**: Bilbao, Valencia
- **Ireland**: Cork
- **UK**: London
- **Greece**: Thessaloniki

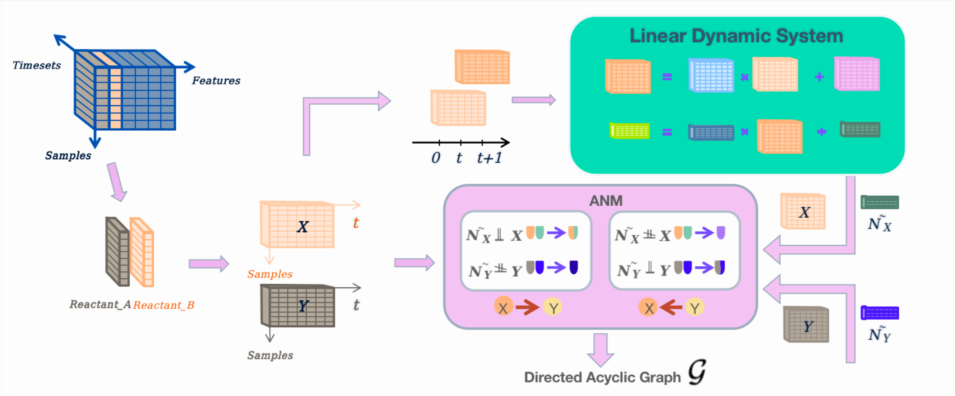
200 Participants



AI Center FEE CTU

Causal learning-IANN

The learning of Nonlinear causal discovery with additive noise models(ANM) can be reduced to the learning of linear dynamic systems.

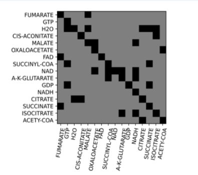


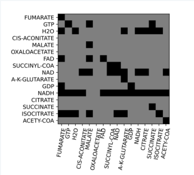
AI Center FEE CTU

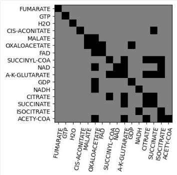
J. Peters, D. Janzing, and B. Schölkopf, "Elements of causal inference: foundations and learning algorithms." *The MIT Press*, 2017.

Future Work

- We considered training a rich class of causal models from time-series data, and we suggest the use of the Krebs cycle and models of metabolism more broadly.
- We established a benchmark for causal learning methodologies within biomedical applications. In this paper, we present a simulated dataset based on the Krebs cycle. Here are adjacency matrices produced by the ground truth and predictions.







Acknowledgements

The research is supervised and supported by Jakub Mareček, PhD, FEE CTU in Prague, Petr Ryšavý, PhD, FEE CTU in Prague and the CoDiet. The CoDiet project is funded by the European Union under Horizon Europe grant number 101084642 and supported by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 101084642].

AI CENTER
FEE CTU

hexiaoyu@fel.cvut.cz

Artificial Intelligence Center
Faculty of Electrical Engineering
Czech Technical University in Prague

Sep 09, 2024