MLCS Journal Club 27.7.21

Latent PCMCI

Moritz Haas

University of Tübingen
**Content**

**Recap: PCMCI**
- 1. Fundamental Idea
- 2. Examples

**Conclusion**
- 1. Advances
- 2. Weaknesses

**LPCMCI**
- 1. Setting
- 2. Idea
- 3. Algorithm
I. Recap PCMCI:
Fundamental Idea and Example
Recap: PCMCI

Goal: learn causal graph of given variables with conditional independence tests

- FullCI has low detection power and estimates low effect size (too many conditionals)
- PC (Peter-Clarke) has high false positive rate for high autocorrelation (broken iid assumption)

PCMCI:
- First PC: Remove irrelevant conditions → learn superset of parents $\hat{\mathcal{P}}(X_j)$ for each variable
- Then MCI (Momentary Conditional Independence):

  \[
  \text{MCI: } X_{t-\tau}^i \perp\!\!\!\perp X_t^j \mid \hat{\mathcal{P}}(X_t^j) \setminus \{X_{t-\tau}^i\}, \hat{\mathcal{P}}(X_{t-\tau}^i).
  \]

  Establish cond. independence \hspace{2cm} Reduce false positive rate

Source: [https://advances.sciencemag.org/content/5/11/eaau4996](https://advances.sciencemag.org/content/5/11/eaau4996)
Start with complete undirected graph

Three steps:
1. Identify the skeleton
2. Identify immoralities and orient them
3. Orient qualifying edges that are incident on colliders

Source: https://www.youtube.com/watch?v=o2A61bJ0UCw
PC Example: Identify the skeleton

\[
A \perp B \mid \{\} \quad \rightarrow \quad X \perp Y \mid \{C\}
\]

\(\forall\) other pairs \((X,Y)\)
PC Example: Orient Edges

\[ A \perp B \mid \emptyset \quad \text{and} \quad X \perp Y \mid \{C\} \quad \forall \text{other pairs } (X,Y) \]

Look at \( X - Y - Z \), where \( X \perp Z \mid S \) and \( Y \) not in \( S \)

We have found all colliders
With conditional independence tests we can only identify the Markov equivalence class.
I. Latent PCMCI:
Setting, Idea and Algorithm
LPCMCI: Setting

**Assumptions:** Faithfulness, Causal Markov Condition, Causal Stationarity

\[ V_t^j = f_j(pa(V_t^j), \eta_t^j) \text{ with } j = 1, \ldots, \tilde{N}. \]

Only observe \( \{X_1, \ldots, X_N\} \subseteq \{V_1, \ldots, V_{\tilde{N}}\} \)

\( \mathcal{M}(\mathcal{G}) \) Maximal Ancestral Graph (MAG) after marginalizing over unobserved variables and all \( t' < t - \tau_{\text{max}} \)

\( \mathcal{P}(\mathcal{G}) \) Maximally informative Partial Ancestral Graph, which can be obtained by running FCI orientation rules (enforce time order and repeating adj.)

May include contemp. links and unobs. confounders
LPCMCI: Idea

Detection power of true link depends on
- Sample size
- CI tests’ sign. level
- CI tests’ estim. dimension
- Effect size $\min_S I(A, B \mid S')$ over used cond. sets $S$

Problem often: low signal-to-noise ratio (e.g. autocorrelation) due to unfortunate choice of conditioning sets leading to low effect size,
Leads to missing links, which lead to false positives and wrong orientations

Carefully choose conditioning sets:
- Discard known non-ancestors
- Include known ancestors ALWAYS
by entangling edge removal and orientation phase

multiple iterations $k$, collecting knowledge
Algorithm 1 LPCMCI

Require: Time series dataset $X = \{X^1, \ldots, X^N\}$, maximal considered time lag $\tau_{\max}$, significance level $\alpha$, CI test $\text{CI}(X, Y, S)$, non-negative integer $k$
1: Initialize $C(\mathcal{G})$ as complete graph with $X_{t-\tau}^i \rightarrow X_t^i$ ($0 < \tau \leq \tau_{\max}$) and $X_{t-\tau}^i \rightarrow X_t^j$ ($\tau = 0$)
2: for $0 \leq l \leq k - 1$ do
3:  Remove edges and apply orientations using Algorithm $S_2$
4:  Repeat line 1, orient edges as $X_{t-\tau}^i \rightarrow X_t^j$ if $X_{t-\tau}^i \rightarrow X_t^j$ was in $C(\mathcal{G})$ after line 3
5: Remove edges and apply orientations using Algorithm $S_2$
6: Remove edges and apply orientations using Algorithm $S_3$
7: return $\text{PAG} C(\mathcal{G}) = \mathcal{P}(\mathcal{G}) = \mathcal{P}(\mathcal{G})_{\tau_{\max}}$

- Introduce node ordering and ‘middle marks’ to save what is known about the link
- There exist several orientation rules for orienting and deleting edges
- Iterate until all middle marks have been removed or reached hyperparam. $k$

In practice for finite samples, larger $k$ does not imply better results :( (see river example)
Figure 2: Results of numerical experiments for (A) LPCMCI(k) for different k, LPCMCI compared to SVAR-FCI and SVAR-RFCI for (B) varying autocorrelation, for (C) number of variables N, and for (D) maximum time lag $\tau_{\text{max}}$ (other parameters indicated in upper right of each panel).
Conclusion: LPCMCI

- Like PCMCI for setting with contemporaneous links and confounders → More general and realistic
- Idea: Always condition on known parents
- Uses middle marks and complicated link orientation and removal rules
- Higher recall for highly autocorr. time series
- order-independent, returns optimal PAG for infinite sample size
- additional hyperparam. k (number of iterations)
- no finite sample guarantees (conv. rate arbitrarily slow for almost unfaithful distr.)


PCMCI: https://advances.sciencemag.org/content/5/11/eaau4996
Thank you!
Important Realisations

- Weak autocorrelation, better recall
- Correct false positive rates in MCI step
- Assumptions: Causal Markov Condition, Faithfulness, causal stationarity, no contemporaneous causal links, no hidden variables
  Then: consistent
- LPCMCI: Causal Markov Condition, Faithfulness, causal stationarity
- Findings of non-causality hold for weaker assumptions (faithfulness, test flexible enough), thus more reliable
- Fast PC: only choose conditions with highest association
- Choose tau_max large, only longer runtime, not higher dim.
- Small alpha means few false positives, but less true positives
- Convergence rate to consistency can be made arbitrarily slow by almost unfaithful distributions
- Autocorrelation breaks iid assumption in tests, thus in ParCorr t distr. Has fewer degrees of freedom hence get more false positives: adjust the degrees of freedom in some way, using pre-whitening, or by block-shuffling. While these approaches help to some extent for the simple bivariate case, they fail in the multivariate case that is relevant for causal discovery [https://arxiv.org/abs/1407.0742](https://arxiv.org/abs/1407.0742) (MCI good option)