

Pot-luck causality challenge: FACT SHEET (for a task solved)

Title: [Results on the PASCAL PROMO challenge](#)

Participant name, address, email and website:

Ivan Markovsky
Building 1, Highfield campus
Southampton, SO17 1BJ, UK

Telephone: +44 (0)23 8059 8715

Fax: +44 (0) 23 8059 4498

Email: im@ecs.soton.ac.uk

WWW: <http://users.ecs.soton.ac.uk/im/homepage.html>

Task(s) solved: [PASCAL PROMO challenge](#)

Reference: <http://eprints.ecs.soton.ac.uk/16779/>

Method:

The data is modeled as a sum of a constant-plus-sin term and a term that is a linear function of a small number of inputs. The problem of identifying such a model from the data is nonconvex in the frequency and phase parameters of the sin and is combinatorial in the number of inputs. The proposed method is suboptimal and exploits several heuristics. First, the problem is split into two phases: 1) identification of the autonomous part and 2) identification of the input dependent part. Second, local optimization method is used to solve the problem in the first phase. Third, ℓ_1 regularization is used in order to find a sparse solution in the second phase.

Results: Please refer to the technical report (<http://eprints.ecs.soton.ac.uk/16779/>) for table with results. In addition, the web page has a link to Matlab software that reproduces the presented results.

Comment about the following:

- quantitative advantages (e.g. compact feature subset, simplicity, computational advantages)

The algorithm is computationally simple: the full model is identified in 3 hours on a standard PC.

- qualitative advantages (e.g. compute posterior probabilities, theoretically motivated, has some elements of novelty).

The tools used to solve the subtasks (leading to the full identification method) are not new however their combination and application for causality detection is novel.

Briefly explain your implementation.

We use Matlab. The one variable nonconvex optimization problem is solved using the Optimization Toolbox (fminsearch function) and the L1 optimization problem is translated to a standard convex optimization problem, using CVX (<http://www.stanford.edu/~boyd/cvx/>).

Provide a URL for the code (if available).

<http://eprints.ecs.soton.ac.uk/16779/2/challenge.tar>

Precise whether it is a push-button application that can be run on benchmark data to reproduce the results, or resources such as modules or libraries.

1. Unpack the archive (it creates a directory called “challenge”).
2. Download and unpack in the same directory the challenge data

<http://www.zurich.ibm.com/~jep/causality/PROMO.zip>

3. If not already installed, download and install CVX

<http://www.stanford.edu/~boyd/cvx/>

4. Make sure that the Optimization Toolbox of Matlab is installed.
5. Change directory to “challenge” and run the function “test” from the Matlab command line. The the model is identified in approximately 3 hours and the results reported in paper (figures and numerical data) are available.

Keywords: Put at *least one keyword in each category*. Try some of the following keywords and add your own:

- Preprocessing or feature construction: redundant input removal.
- Causal discovery: prediction, least squares fitting.
- Feature selection: **L1 norm regularization**.