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

Title: ***Iterative Stepwise Selection and Threshold for Learning Causes in Simple Time Series***

Participant name, address, email and website:
Jianxin Yin **jianxinyin@math.pku.edu.cn**
Shaopeng Wang **wangshop@gmail.com**
Wanlu Deng **shirleypku@gmail.com**
Ya Hu **teresahu@pku.edu.cn**
Zhi Geng **zgeng@math.pku.edu.cn**
School of Mathematical Sciences
Peking University
Beijing 100871, China

Task(s) solved: **PROMO--Simple causal effects in time series**

Reference:
Provide a pointer to a technical memorandum or a paper (optional).
Yin et al., ***Iterative Stepwise Selection and Threshold for Learning Causes in Simple Time Series***

Method:
Summarize the algorithms you used in a way that those skilled in the art should understand what to do. Profile of your methods as follows:

- Preprocessing
  We use several vectors of function *sin*, *cos* with different periods and the same length (1095) to represent a "pseudo" design matrix. Then we get the regression estimate for the products and promotions and find the regression residuals for products and promotions. This process is considered as the seasonal removing step. The residues obtained at this step will be analyzed at the next step.
- Causal discovery & Feature selection
  Here we assume that promotions are the potential causes for the product sales and not the reverse and that the causal relationships between sales and promotions have linear structure equation models. So the causal discovery problem is reduced to the problem of variable selection in a linear regression framework after removing the seasonal term. We use the stepwise approach for model selection, in which we use an iterative process until there is no change of selected sets between two consecutive iterations.

We use the hard threshold to select significant influential promotions from the output of the iterative stepwise selection. The threshold is performed on the regression coefficients of the normalized covariates. All the coefficients of the promotions in the current model are sorted in their absolute values. For a fixed ratio a ($0<a<0.5$), we drop those coefficients whose absolute values are bellow a*100% of the largest absolute value. Then we repeat this process iteratively until the selected variable set does not change. In every iteration, we build a regression model on the remaining set of the previous iteration and then use the fixed a to threshold the current coefficients as described above.

- <u>Model selection/hyperparameter selection</u>
  We select the hyper-parameter a using extended-BIC criterion.

Results: The reader should also know from reading the fact sheet what the strength of the method is. To that end, provide a result table:

<u>Table 1: Result table.</u>

Comment about the following:

- <u>quantitative advantages</u> (e.g. compact feature subset, simplicity, computational advantages)
  Our algorithm has a fast speed in computation and simple enough in model building.
- <u>qualitative advantages</u> (e.g. compute posterior probabilities, theoretically motivated, has some elements of novelty).
  The iterative applying of the stepwise selection is not common in our usual applications. And the iterative thresholding on the regression coefficients as a method of feature selection is also novel. Under certain assumptions (the signal-to-noise ratio is above a bound), our method is robust to different types of the underlying models for the time series (e.g. AR, MA, ARMA). This is partly due to the fitting of the seasonal term by using different periods of *sin-cos* sequences. For the Fourier series can approximate a continuous function as accurate as possible. And the underlying seasonal term is a continuous function of time as our basic assumption.

Briefly explain your implementation. Provide a URL for the code (if available). 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.

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

- <u>Preprocessing or feature construction</u>: linear regression, sin-cos functions, “pseudo” design matrix.
- <u>Causal discovery&Feature selection</u>: iterative, stepwise selection, hard threshold.
- <u>Hyper-parameter selection</u>: extended-BIC criterion