Causal Dynamic Time Lag: Predicting What & When

M. Chandorkar, C. Furthlener, E. Camporeale and M. Sebag
CWI Amsterdam, INRIA Paris Saclay

Abstract

We formalize the joint regression task of predicting the magnitude of signals as well as the time delay with respect to their driving phenomena.

1 Introduction

2 Causality in Time Series

First level headings are all caps, flush left, bold, and in point size 12. Use one line space before the first level heading and one-half line space after the first level heading.

2.1 Granger Causality

Second level headings are initial caps, flush left, bold, and in point size 10. Use one line space before the second level heading and one-half line space after the second level heading.

2.1.1 Previous Work

Third level headings are flush left, initial caps, bold, and in point size 10. Use one line space before the third level heading and one-half line space after the third level heading.

2.2 CITATIONS, FIGURES, REFERENCES

2.2.1 Citations in Text

Citations within the text should include the author's last name and year, e.g., (Cheesman, 1985). References should follow any style that you are used to using, as long as their style is consistent throughout the paper. Be sure that the sentence reads correctly if the citation is deleted: e.g., instead of "As described by (Cheesman, 1985), we first frobulate the widgets," write "As described by Cheesman (1985), we first frobulate the widgets."

2.2.2 Footnotes

Indicate footnotes with a number[1] in the text. Use 8 point type for footnotes. Place the footnotes at the bottom of the column in which their markers appear, continuing to the next column if required. Precede the footnote section of a column with a 0.5 point horizontal rule 1 inch (6 picas) long.[2]

2.2.3 Figures

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction, and art work should not be hand-drawn. Figures may appear at the top of a column, at the top of a page spanning multiple columns, inline within a column, or with text wrapped around them, but the figure number and caption always appear immediately below the figure. Leave 2 line spaces between the figure and the caption. The figure caption is initial caps and each figure should be numbered consecutively.

Make sure that the figure caption does not get separated from the figure. Leave extra white space at the bottom of the page rather than splitting the figure and figure caption.

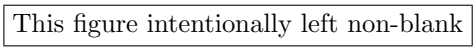

Figure 1: Sample Figure Caption

Preliminary work. Under review by AISTATS 2019. Do not distribute.

[1] Sample of the first footnote.
[2] Sample of the second footnote.

2.2.4 Tables

All tables must be centered, neat, clean, and legible. Do not use hand-drawn tables. Table number and title always appear above the table. See Table 1.

Use one line space before the table title, one line space after the table title, and one line space after the table. The table title must be initial caps and each table numbered consecutively.

Table 1: Sample Table Title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

3 Problem Formulation

Causal Dynamic Time Lag is essentially a regression problem with two tasks. Given two time series, the causes $x(t)$ and the observed effects $y(t)$, the regression model must learn a mapping $f()$ which maps each input patten $x(t_1)$ to an output $y(t_2)$, and a mapping $g()$ which maps the time delay between the input and output patterns $t_2 = t_1 + g(x(t_1))$. This is formally specified in equations below.

$$t \in \mathbb{R}^+ \tag{1}$$
$$x(t) \in \mathcal{X} \tag{2}$$
$$f : \mathcal{X} \to \mathbb{R} \tag{3}$$
$$g : \mathcal{X} \to \mathbb{R}^+ \tag{4}$$
$$\Delta(t) = g[x(t)] \tag{5}$$
$$y(t + \Delta(t)) = f[x(t)] \tag{6}$$

4 Proposed Solution

In practical time-series applications, one works with sub-sampled or discretized versions of the time series $x(t)$ and $y(t)$. The time delay function $g(.)$ can now be recast as a function which for every input pattern $x(t_i)$, returns a time delay Δ corresponding the the time step $i + \Delta$ when the effect $y(t_i + \Delta)$ is observed.

For practical purposes one must define for every time step t, a *causal time window* $[t + \ell, t + \ell + h]$, within which the model searches for probably temporal causal links.

Our proposed model then produces the following predictions:

1. Targets $y(t + \ell), \cdots, y(t + \ell + h - 1)$
2. Time Lag Probabilities $\hat{p}(t+\ell), \cdots, \hat{p}(t+\ell+h-1)$

The model thus tries to learn a predictor for each lagged output $y(t + i)$ in the causal window $[t + \ell, t + \ell + h]$, and simultaneosly supplies a probability $\hat{p}(t+i)$ which represents the likelihood of a causal link between $x(t)$ and $y(t + i)$.

5 Loss Function

Balance two incentives

1. Generate accurate predictions for time window $y(t + \ell), \cdots, y(t + \ell + h - 1)$
2. Learn time lag structure according to some intuition.

$$\mathcal{L}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)}) = \lambda_1 \sum_{i,m} \frac{1}{2M} (y_i^{(m)} - \hat{y}_i^{(m)})^2 (1 + \hat{p}_i^{(m)}) + \lambda_2 \mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)})$$

The term $\mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)})$ penalizes the predicted probabilities $\hat{p}^{(1:M)}$, for deviation from some chosen *target probability*.

From the intuitions of Granger causality, we use the concept of causality as predictability, we can thus characterize the *target probability*, $\widetilde{p}$ for a time window $[t + \ell, t + \ell + h - 1]$ in the following manner: The lagged output $y(t+i)$ which has greater predictability given $x(t)$, is a more likely causal link.

$$\mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)}) = \sum_{m=1}^{M} \frac{1}{M} \mathcal{H}(\hat{p}^{(m)}, \widetilde{p}^{(m)}) \tag{7}$$
$$\mathcal{H}(p, q) = \sqrt{\sum_i (\sqrt{p_i} - \sqrt{q_i})^2} \tag{8}$$
$$\widetilde{p}^{(m)} = softmax\left(\frac{1}{T}(y^{(m)} - \hat{y}^{(m)})^2\right) \tag{9}$$

6 Experiments

For the camera-ready paper, if you are using LaTeX, please make sure that you follow these instructions. (If you are not using LaTeX, please make sure to achieve the same effect using your chosen typesetting package.)

1. Download `fancyhdr.sty` – the `aistats2019.sty` file will make use of it.

2. Begin your document with

   ```
   \documentclass[twoside]{article}
   \usepackage[accepted]{aistats2019}
   ```

   The `twoside` option for the class article allows the package `fancyhdr.sty` to include headings for even and odd numbered pages. The option `accepted` for the package `aistats2019.sty` will write a copyright notice at the end of the first column of the first page. This option will also print headings for the paper. For the *even* pages, the title of the paper will be used as heading and for *odd* pages the author names will be used as heading. If the title of the paper is too long or the number of authors is too large, the style will print a warning message as heading. If this happens additional commands can be used to place as headings shorter versions of the title and the author names. This is explained in the next point.

3. If you get warning messages as described above, then immediately after `\begin{document}`, write

   ```
   \runningtitle{Provide here an alternative
   shorter version of the title of your
   paper}
   \runningauthor{Provide here the surnames
   of the authors of your paper, all
   separated by commas}
   ```

   Note that the text that appears as argument in `\runningtitle` will be printed as a heading in the *even* pages. The text that appears as argument in `\runningauthor` will be printed as a heading in the *odd* pages. If even the author surnames do not fit, it is acceptable to give a subset of author names followed by "et al."

4. Use the file sample_paper.tex as an example.

5. The camera-ready versions of the accepted papers are 8 pages, plus any additional pages needed for references.

6. If you need to include additional appendices, you can include them in the supplementary material file.

7. Please, don't change the layout given by the above instructions and by the style file.

Acknowledgements

Use the unnumbered third level heading for the acknowledgements. All acknowledgements go at the end of the paper.

References

References follow the acknowledgements. Use an unnumbered third level heading for the references section. Any choice of citation style is acceptable as long as you are consistent. Please use the same font size for references as for the body of the paper—remember that references do not count against your page length total.

References

J. Alspector, B. Gupta, and R. B. Allen (1989). Performance of a stochastic learning microchip. In D. S. Touretzky (ed.), *Advances in Neural Information Processing Systems 1*, 748–760. San Mateo, Calif.: Morgan Kaufmann.

F. Rosenblatt (1962). *Principles of Neurodynamics.* Washington, D.C.: Spartan Books.

G. Tesauro (1989). Neurogammon wins computer Olympiad. *Neural Computation* **1**(3):321–323.