Non-parametric estimation and inference for conditional mass based Granger causality measures for high-dimensional Markov chains

Xiaoxiao Ma* Kaveh Salehzadeh Nobari†

University of Leeds Durham University

September 16, 2020

This is a preliminary draft. Please do not cite or distribute without the permission of the authors.

*Institute for Transport Studies (ITS), University of Leeds. Address and email: Institute for Transport Studies (ITS), University of Leeds, 34-40 University Road, Leeds, LS2 9JT

†Department of Economics and Finance, Durham University. Address and email: Department of Economics and Finance, Durham University Business School, Mill Hill Lane, Durham, DH1 3LB.

Non-parametric estimation and inference for conditional mass based Granger causality measures for high-dimensional Markov chains

ABSTRACT

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Journal of Economic Classification: C12, C13, C14

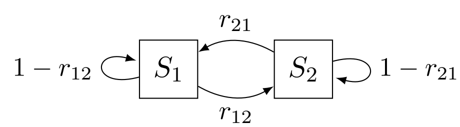
Keywords: Causality measures; Non-parametric estimation; Vine decomposition; Markov chains; Bernstein copula density; Local bootstrap.

Contents

1 Introduction

2 Motivation

Consider the simple case of a two-state Markov chain, such as the one depicted below



where

$$r_{ij} = P[X_t = i \mid X_{t-1} = j], \quad i, j \in \{1, 2\} \tag{1}$$

To estimate the conditional probabilties r_{ij}, we may express (1) as follows

$$r_{ij} = \frac{P[X_t = i, X_{t-1} = j]}{P[X_{t-1} = j]} \tag{2}$$

We know from the Theorem of Sklar (1959) that for the continuous variables $\{(X, Y, Z) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R} \equiv \mathbb{R}^3\}$, the joint distribution function $F_{XYZ}(x, y, z)$ can be expressed using the marginal distributions and a copula function capturing the depedence, - i.e.

$$F(x, y, z) = C(F_X(x), F_Y(y), F_Z(z)) \tag{3}$$

where $F_\kappa(.)$ for $\kappa = X, Y, Z$ is the marginal distribution function of variable κ and $C(F_X(.), F_Y(.), F_Z(.))$ is a copula function on $[0.1]^3$ that captures the dependence of (X, Y, Z). If we take the derivative of 3 with respect to (x, y, z), we obtain the joint density funciton of (X, Y, Z):

$$f(x, y, z) = f_X(x) \times f_Y(y) \times f_Z(z) \times c(F_X(x), F_Y(y), F_Z(z)) \tag{4}$$

3 Copula-based Granger causality measures for qualitative processes

Consider a homogeneous Markov process of order one with a state-space consisting of finite elements and the transition matrix

$$P[X_t = i, Y_t = k \mid X_{t-1} = j, Y_{t-1} = l] = p_{i,k|j,l}, \tag{5}$$

for $i, j = 1, \cdots, J$ and $k, l = 1, \cdots, L$. Furthermore, the invariant initial probability is given by

$$P[X_t = i, Y_t = k] = \pi_{i,k} \tag{6}$$

with $i = 1, \cdots, J$ and $k = 1, \cdots, L$. Gourieroux et al. (1987) show that the causality measure from X to Y is defined by

$$C(X \to Y) = \sum_{j=1}^{J} \sum_{l=1}^{L} \pi_{j,l} \left\{ \sum_{k=1}^{L} p_{k|j,l} \log \frac{p_{k|j,l}}{p^{y}_{k|l}} \right\} \tag{7}$$

where

$$p_{k|j,l} = \sum_{i=1}^{J} p_{i,j|k,l} = P[X_t = i \mid X_{t-1} = j, Y_{t-1} = l] \quad \text{and} \quad p^{y}_{k|l} = \frac{\sum_{k=1}^{L} \sum_{l=1}^{L} p_{i,k|j,l} \pi_{j,l}}{\sum_{l=1}^{L} \pi_{j,l}}$$

However, expression (7) is difficult to evaluate.

$$P[X_t = i, Y_t = k \mid X_{t-1} = j, Y_{t-1} = l] = p_{i,k|j,l}$$

$$f[x^*_t, y^*_t \mid x^*_{t-1}, y^*_{t-1}] = p_{i,k|j,l}$$

with

$$X^*_t = X_t + U - 1, \quad U \sim Uniform[0, 1]$$

$$f[x^*_t, y^*_t, x^*_{t-1}, y^*_{t-1}]/f(x^*_{t-1}, y^*_{t-1}) =$$
$$f(x^*_t)f(y^*_t)f(x^*_{t-1})f(y^*_{t-1})c(F(x^*_t), F(y^*_t), F(x^*_{t-1}), F(y^*_{t-1}))/f(x^*_{t-1})f(y^*_{t-1})c(F(x^*_{t-1}), F(y^*_{t-1})$$

$$f[y^*_t, x^*_{t-1}, y^*_{t-1}] = f(y^*_t)f(x^*_{t-1})f(y^*_{t-1})c(F(y^*_t), F(1), F(y^*_{t-1}), F(x_{t-1}))$$

4 Estimation and inference

4.1 Estimation

4.2 Inference

5 Measuring causality in high-dimensional Markov chains

6 Monte Carlo simulations

6.1 Bootstrap bias-corrected estimator of Granger causality measures

6.1.1 Bootstrap bias-correction

6.1.2 Simulation study

6.2 Empirical size and power

7 Empirical application

8 Conclusion

References

Gourieroux, C., Monfort, A., and Renault, E. (1987). Kullback causality measures. *Annales d'Economie et de Statistique*, pages 369–410.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231.

9 Appendix: Proofs