GenMarkov: Modeling Generalized Multivariate Markov Chains in R

by C. Vasconcelos, B. Damásio

Supplementary Materials

Consider the following proposed model:

$$P(S_{jt} = i_0|S_{1t-1} = i_1, \dots, S_{st-1} = i_s, \boldsymbol{x}_t) \equiv \\ \lambda_{j1} P(S_{jt} = i_0|S_{1t-1} = i_1, \boldsymbol{x}_t) + \dots + \lambda_{js} P(S_{jt} = i_0|S_{st-1} = i_s, \boldsymbol{x}_t) \quad (1)$$

s.t. $\sum_{k=1}^{s} \lambda_{jk}$ and $\lambda_{jk} \geq 0$.

The following Proposition holds the consistency of the MLE estimator:

Proposition 1. *Let $\{\boldsymbol{w}_t\}$ be ergodic stationary random variable, with likelihood function $f(\boldsymbol{w}_t|\theta_0)$. Let $\widehat{\theta}$ be the MLE estimator. Suppose that:*

1. $E[\log f(\boldsymbol{w}_t \mid \theta_0)]$ *is uniquely maximized on Θ at $\theta_0 \in \Theta$,*
2. $\theta_0 \in \Theta$*, which is compact,*
3. $\log f(\boldsymbol{w}_t \mid \theta_0)$ *is continuous at each $\theta \in \Theta$ with probability one,*
4. $E[sup_{\theta \in \Theta} \mid \log f(w_t \mid \theta) \mid] < \infty$

Then $\widehat{\theta} \xrightarrow{p} \theta_0$

Condition (1) is verified according to Lemma 2.2 of Newey and Mcfadden (1994). Condition (2) is verified and guaranteed by the restrictions imposed in the model parameters. Knowing that $P(S_{jt} = i_0|S_{1t-1} = i_1, \dots, S_{st-1} = i_s, \boldsymbol{x}_t)$ is linear combination of a set of n probabilities and since the logarithm function is a continuous function, condition (3) is verified. Finally, condition (4) is verified according to Lemma 2.4 of Newey and Mcfadden (1994).

Regarding inference, MLE will be asymptotically normal if it is consistent and the following Proposition verifies:

Proposition 2. *Let $\{\boldsymbol{w}_t\}$ be ergodic stationary random variable and let $s(\boldsymbol{w}_t;\theta)$ and $H(\boldsymbol{w}_t;\theta)$ be the first and second partial derivatives of the $\log f(\boldsymbol{w}_t \mid \theta)$, respectively. Suppose the estimator $\widehat{\theta}$ is consistent and suppose, further, that*

1. *θ_0 is in the interior of Θ,*
2. *$\log f(\boldsymbol{w}_t \mid \theta_0)$ is twice continuously differentiable in θ for any $\boldsymbol{w}_t$,*
3. *$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} s(\boldsymbol{w}_t;\theta_0) \underset{d}{\to} N(0,\Sigma)$, where Σ is positive definite,*
4. *For some neighborhood of $\mathcal{N}$ of θ_0,*

$$E[\sup_{\theta\in\mathcal{N}} ||H(\boldsymbol{w}_t;\theta)||] < \infty$$

so that for any consistent estimator $\tilde{\theta}$, $\frac{1}{n}\sum_{t=1}^{n} H(\boldsymbol{w}_t;\tilde{\theta}) \underset{p}{\to} E[H(\boldsymbol{w}_t;\theta)]$ v. $E[H(\boldsymbol{w}_t,\theta_0)]$ is nonsingular.

Then $\widehat{\theta}$ is asymptotically normal with

$$Avar(\widehat{\theta}) = \{E[H(\boldsymbol{w}_t;\theta_0)]\}^{-1}\Sigma\{E[H(\boldsymbol{w}_t;\theta_0)]\}^{-1}$$

Condition (1) is verified as in condition (2) of Proposition (1). Condition (2) is also verified as in condition (3) of Proposition (1), since the logarithm function is twice continuously differentiable. Condition (3) is verified according to the Ergodic Stationary Martingale Differences CLT Billingsley (1961). In this case, $\Sigma = E[s(\boldsymbol{w}_t;\theta_0)s(\boldsymbol{w}_t;\theta_0)'] = -E[H(\boldsymbol{w}_t;\theta_0)]$, which implies that $Avar(\widehat{\theta}) = -\{E[H(\boldsymbol{w}_t;\theta_0)]\}^{-1}$. Condition (4) is verified according to Lemma 2.4 of Newey and Mcfadden (1994). Considering only one equation, let $\boldsymbol{q}_t$ be a $t \times s$ matrix of the probabilities $P(S_{1t} \mid S_{1t}, x_t), \ldots, P(S_{1t} \mid S_{st}, x_t)$ and $\boldsymbol{\lambda}$ a row-vector of $\lambda_{11}, \ldots, \lambda_{1s}$, the hessian matrix is given by $E[\boldsymbol{q}_t'\boldsymbol{q}_t[(\boldsymbol{\lambda}\boldsymbol{q}_t')(\boldsymbol{\lambda}\boldsymbol{q}_t')']^{-1}]$. Condition (v) is verified if $E[\boldsymbol{q}_t'\boldsymbol{q}_t]$ is nonsingular.

Bibliography

P. Billingsley. The lindeberg-lévy theorem for martingales. *Proceedings of the American Mathematical Society*, 12(5):788–792, 1961. URL http://www.jstor.org/stable/2034876. [p2]

W. Newey and D. Mcfadden. Chapter 36 large sample estimation and hypothesis testing. In *Handbook of Econometrics*, volume 4, pages 2111–2245. Elsevier, 1994. doi: https://doi.org/10.1016/S1573-4412(05)80005-4. [p1, 2]

Carolina Vasconcelos
NOVA Information Management School (NOVA IMS)
Campus de Campolide, 1070-312 Lisboa, Portugal
cvasconcelos@novaims.unl.pt

Bruno Damásio
NOVA Information Management School (NOVA IMS)
Campus de Campolide, 1070-312 Lisboa, Portugal
bdamasio@novaims.unl.pt