Conditional Independence between Markov Chains (Simulations) [1]

Márcio Luis Lanfredi Viola

Jesus Enrique Garcia

Verónica Andrea González-López

Instituto de Matemática Estatística e Computação Científica - IMECC
Universidade Estadual de Campinas - UNICAMP

1 Criterion of the conditional independence

The aim of this work is to develop a methodology for investigating the conditional dependence between Markov chains, that is, the objective is to obtain a procedure to verify if the process are conditionally dependents. This procedure is based on the Bayesian information criterion (BIC) [1, 2].

Considere the finite order Markov stationary process, $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$, with same alphabet $\mathcal{A}$. Moreover, considere that $(Z_t)_{t\in\mathbb{Z}} = (X_t, Y_t)_{t\in\mathbb{Z}}$ is an finite order Markov stationary process. Considere $r \in \mathcal{A}^k$ and $s \in \mathcal{A}^k$ fixed strings.

Definição 1. *Considere $r \in \mathcal{A}^k$ and $s \in \mathcal{A}^k$ fixed strings. The k order Markov stationary process, $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$, are conditionally independents given (r, s) if*

$$\begin{aligned} Prob(X_{t+1} = a \quad , \quad & Y_{t+1} = b | X_{t-k+1}^{t} = r, Y_{t-k+1}^{t} = s) = \\ & Prob(X_{t+1} = a | X_{t-k+1}^{t} = r) Prob(Y_{t+1} = b | Y_{t-k+1}^{t} = s) \end{aligned}$$

for any $a, b \in \mathcal{A}$ and for any t.

Let o_{Z^1} the order of the process $(Z_t)_{t\in\mathbb{Z}}$ when $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ aren't conditionally independents given (r, s) and o_{Z^O} the order of the process $(Z_t)_{t\in\mathbb{Z}}$ when $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are conditionally independents given (r, s). Considere $k = \max\{o_{Z^1}, o_{Z^O}\}$.

[1] This work is supported by CAPES

Let $x_1^n = x_1...x_{n-1}...x_n$, $y_1^n = y_1...y_{n-1}y_n$ e $z_1^n = (x_1, y_1)...(x_n, y_n)$ samples, respectively, of the process $(X_t)_{t\in\mathbb{Z}}$, $(Y_t)_{t\in\mathbb{Z}}$ and $(Z_t)_{t\in\mathbb{Z}}$.

The BIC of process $(Z_t)_{t\in\mathbb{Z}}$, when the process $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are conditionally independents, is

$$\begin{aligned} BIC^o(z_1^n) = & - \sum_{i=1}^{k}\sum_{a=1}^{k} N^X(i,a)\log\left(\frac{N^X(i,a)}{N^X(i)}\right) - \sum_{j=1}^{k}\sum_{b=1}^{k} N^Y(j,b)\log\left(\frac{N^Y(j,b)}{N^Y(j)}\right) \\ & + \frac{(|\mathcal{A}|-1)|\mathcal{A}|^k}{2}\log(n) + \frac{(|\mathcal{A}|-1)|\mathcal{A}|^k}{2}\log(n) \end{aligned} \tag{1.1}$$

where $N^X(i,a)$ is the number of occurrences of the string $i \in \mathcal{A}^k$ followed by the letter $a \in \mathcal{A}$ in the sample x_1^n and $N^Y(j,b)$ is the number of occurrences of the string $j \in \mathcal{A}^k$ followed by the letter $b \in \mathcal{A}$ in the sample y_1^n, such that, $N^X(i,a) = \left|\{k < m \leq n : x_{m-k}^{m-1} = i, x_m = a\}\right|$ and $N^Y(j,b) = \left|\{k < m \leq n : y_{m-k}^{m-1} = j, y_m = b\}\right|$.

The BIC of process $(Z_t)_{t\in\mathbb{Z}}$, when the process $(X_t)_{t\in\mathbb{Z}}$ and $(Y_t)_{t\in\mathbb{Z}}$ are conditionally dependents given (r,s), is

$$\begin{aligned} BIC^o((r,s),z_1^n) = & - \sum_{\substack{i,j\in\mathcal{A}^k \\ a,b\in\mathcal{A} \\ i\neq r \\ j\neq s}} N^{XY}((i,a),(j,b))\log\left(\frac{N^{XY}((i,a),(j,b))}{N^{XY}(i,j)}\right) \\ & - \sum_{a,b\in\mathcal{A}} N^{XY}((r,a),(s,b))\log\left(\frac{N^X(r,a)}{N^X(r)}\right) - \sum_{a,b\in\mathcal{A}} N^{XY}((r,a),(s,b))\log\left(\frac{N^Y(s,b)}{N^Y(s)}\right) \\ & + \frac{(|\mathcal{A}|^2-1)(|\mathcal{A}|^{2k}-1)}{2}\log(n) + \frac{|\mathcal{A}|-1}{2}\log(n) + \frac{|\mathcal{A}|-1}{2}\log(n) \end{aligned} \tag{1.2}$$

where $N^{XY}((i,a),(j,b))$ is the number of occurrences of the string $i \in \mathcal{A}^k$ followed by the letter $a \in \mathcal{A}$ in the sample x_1^n and of the string $j \in \mathcal{A}^k$ followed by the letter $b \in \mathcal{A}$ in the sample y_1^n, such that, $N^{XY}((i,a),(j,b)) = \left|\{k < m \leq n : (x_{m-k}^{m-1} = i, y_{m-k}^{m-1} = j), (x_m = a, y_m = b)\}\right|$

The propopsted procedure establishes that

$$\lim_{n\to\infty} I_{\left\{BIC^1(z_1^n) - BIC^o((r,s),z_1^n)\geq 0\right\}} = 1$$

if and only if

$$P^{XY}_{(ra)(sb)} = P^X_{ra}P^Y_{sb}$$

for any $a, b \in \mathcal{A}$ and $i, j \in \mathcal{A}^k$ where $P^X_{ra} = Prob(X_{t+1} = a|X^t_{t-k} = r)$, $P^Y_{sb} = Prob(Y_{t+1} = b|Y^t_{t-k} = s)$, $P^{XY}_{(ra)(sb)} = Prob((X_{t+1} = a, Y_{t+1} = b)|(X^t_{t-k} = r, Y^t_{t-k} = s))$, $P^{XY}_{(ra)(sb)} = Prob(X_{t+1} = a, Y_{t+1} = b|X^t_{t-k} = r, Y^t_{t-k} = s)$ and $I(.)$ is the characteristic function.

We show the proof of the relation above and show three simulation scenarios.

2 Simulations

2.1 Simulation 1:

Considere the Markov stationary process of order 1, $(Z_t)_{t\in\mathbb{Z}} = (X_t, Y_t)_{t\in\mathbb{Z}}$, with alphabet $\{0, 1, 2, 3\}$. The transition probabilities $P^{XY}_{(ra)(sb)}, \forall a, b, r, s \in \{0, 1, 2, 3\}$ are shown by matrix (2.3) where their rows and columns represent, respectively, r (ou s) e a (ou b).

$$T^Z = \begin{pmatrix} 0,25 & 0,25 & 0,25 & 0,25 \\ 0,2 & 0,3 & 0,2 & 0,3 \\ 0,2 & 0,2 & 0,3 & 0,3 \\ 0,16 & 0,24 & 0,24 & 0,36 \end{pmatrix} \tag{2.3}$$

We simulated 500 samples, the length 3.000, 10.000, 15.000 e 50.000, of the process $(Z_t)_{t\in\mathbb{Z}}$.

2.2 Simulation 2:

Considere the Markov stationary process of order 2, $(Z_t)_{t\in\mathbb{Z}} = (X_t, Y_t)_{t\in\mathbb{Z}}$, with alphabet $\{0, 1, 2, 3\}$. The transition probabilities $P^{XY}_{(ra)(sb)}, \forall a, b \in \{0, 1, 2, 3\}$,

$\forall r, s \in \{00, 10, 20, 30, 01, 11, 21, 31, 02, 12, 22, 32, 03, 13, 23, 33\}$, are shown by matrix (2.4).

$$T^Z = \begin{pmatrix} 0,36 & 0,24 & 0,24 & 0,16 \\ 0,18 & 0,42 & 0,12 & 0,28 \\ 0,18 & 0,12 & 0,42 & 0,28 \\ 0,09 & 0,21 & 0,21 & 0,49 \\ 0,18 & 0,42 & 0,12 & 0,28 \\ 0,43 & 0,18 & 0,28 & 0,11 \\ 0,09 & 0,21 & 0,21 & 0,49 \\ 0,21 & 0,09 & 0,49 & 0,21 \\ 0,24 & 0,16 & 0,36 & 0,24 \\ 0,12 & 0,28 & 0,18 & 0,42 \\ 0,48 & 0,32 & 0,12 & 0,08 \\ 0,24 & 0,56 & 0,06 & 0,14 \\ 0,12 & 0,28 & 0,18 & 0,42 \\ 0,28 & 0,12 & 0,42 & 0,18 \\ 0,24 & 0,56 & 0,06 & 0,14 \\ 0,56 & 0,24 & 0,14 & 0,06 \end{pmatrix} \tag{2.4}$$

We simulated 500 samples, the length 3.000, 10.000, 15.000 e 50.000, of the process $(Z_t)_{t\in\mathbb{Z}}$.

2.3 Simulation 3:

Considere the Markov stationary process of order 1, $(Z_t)_{t\in\mathbb{Z}} = (X_t, Y_t)_{t\in\mathbb{Z}}$, with alphabet $\{0, 1, 2, 3\}$. The transition probabilities $P^{XY}_{(ra)(sb)}, \forall a, b, r, s \in \{0, 1, 2, 3\}$, are shown by matrix (2.5).

$$T^Z = \begin{pmatrix} 0,25 & 0,25 & 0,25 & 0,25 \\ 0,5 & 0 & 0 & 0,5 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,5 & 0 & 0 & 0,5 \end{pmatrix} \tag{2.5}$$

We simulated 500 samples, the length 3.000, 10.000, 15.000 e 50.000, of the process $(Z_t)_{t\in\mathbb{Z}}$.

Referências

[1] Csiszár, I., and Talata, Z. Context tree estimation for not necessarily finite memory processes, via bic and mdl. *IEEE Transactions on Information Theory 52*, 3 (2006), 1007–1016.

[2] García, J. E., and González-López, V. A. Minimal markov models. *Can be retrieved from arXiv:1002.0729v1* (2010).