Trabalho Nº6 - Modelos Markovianos

Rafael Morciani | GRR20160217

14 de Setembro de 2020

Exercicio 1a

1. Considere $\{S_t, C_t \ : \ t \in \mathbb{N}\}$ um Modelo Oculto de Markov Poisson com dois estados, de parâmetros:

$$\Gamma = \begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix} \qquad \text{e} \qquad \lambda = (1, 3)\cdot$$

Em cada uma das seguintes maneiras, calcule a probabilidade de que as três primeiras observações desse modelo sejam 0, 2, 1.

(a) Considere todas as possíveis sequências de estados da Cadeia de Markov que poderiam ter ocorrido. Calcule a probabilidade de cada sequência e a probabilidade das observações dadas a cada sequência.

Desenvolvimento A

Da propriedade de Markov e da estacionariedade da sequência C_t onde t = 1, 2, 3, temos:

$$P(C_1 = i, C_2 = j, C_3 = k) = P(C_1 = i).P(C_2 = j|C_1 = j).P(C_3 = k|C_2 = j) = \delta_i \gamma_{ij} \gamma_{jk}$$

Usando a propriedade de independência condicional, obtemos:

$$P(S_1 = 0, S_2 = 2, S_3 = 1) =$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} P(S_1 = 0, S_2 = 2, S_3 = 1|C_1 = i, C_2 = j, C_3 = k)P(C_1 = i, C_2 = j, C_3 = k)$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} p_i(0)p_j(2)p_k(1)\delta_i \gamma_{ij} \gamma_{jk}, \ Onde \ p_l(s) = \frac{\lambda_l^s e^{-\lambda_l}}{s!}, \ l \, \epsilon \, \{1, 2\}$$

Então temos:

$$(\delta_1\ \delta_2) = \frac{1}{0.9+0.4}(0.4\ 0.9) = (4/13\ 9/13)$$

E com o valor dos lambdas $\lambda_1 = 1\ e\ \lambda_2 = 3$ obtemos:

$$p_1(0) \approx 0,368\ ,\ p_1(2) \approx 0,184\ ,\ p_1(1) \approx 0.368$$

$$p_2(0) \approx 0,05\ ,\ p_2(2) \approx 0,224\ ,\ p_2(1) \approx 0,149$$

A tabela a seguir lista todas as sequências possíveis de estados e as respectivas probabilidades:

i	j	k	$p_i(0)$	$p_j(2)$	$p_k(1)$	δ_i	γ_{ij}	γ_{jk}	Π
1	1	1	0.368	0.184	0.368	4/13	0.1	0.1	0.000077
1	1	2	0.368	0.184	0.149	4/13	0.1	0.9	0.000280
1	2	1	0.368	0.224	0.368	4/13	0.9	0.4	0.003359
1	2	2	0.368	0.224	0.149	4/13	0.9	0.6	0.002045
2	1	1	0.050	0.184	0.368	9/13	0.4	0.1	0.000093
2	1	2	0.050	0.184	0.149	9/13	0.4	0.9	0.000341
2	2	1	0.050	0.224	0.368	9/13	0.6	0.4	0.000682
2	2	2	0.050	0.224	0.149	9/13	0.6	0.6	0.000415
								Σ	0.007292

Observe que os valores da tabela podem ser usados para calcular a probabilidade condicional de cada sequência possível de estados, C_1, C_2, C_3 usando:

$P(C_1 = i, C_2 = j, C_3 = k | S_1 = 0, S_2 = 2, S_3 = 1)$

$= \frac{P(C_1=i, C_2=j, C_3=k, S_1=0, S_2=2, S_3=1)}{P(S_1=0, S_2=2, S_3=1)},\ para\ i, j, k\ \epsilon\ \{1, 2\}$

Assim, a sequência de estados mais provável aqui é $C_1 = 1, C_2 = 2, C_3 = 1$ sua probabilidade condicional é dada por $\frac{0.003359}{0.007292} \approx 0.461$.

Exercicio 1b

(b) Aplique a expressão

$$P(S_1 = 0, S_2 = 2, S_3 = 1) = \delta P(0)\Gamma P(2)\Gamma P(1)\mathbf{1}^\top,$$

onde

$$P(s) = \begin{pmatrix} \frac{\lambda_1^s}{s!}e^{-\lambda_1} & 0 \\ 0 & \frac{\lambda_2^s}{s!}e^{-\lambda_2} \end{pmatrix} = \begin{pmatrix} \frac{1^s}{s!}e^{-1} & 0 \\ 0 & \frac{3^s}{s!}e^{-3} \end{pmatrix}.$$

Desenvolvimento B

$$P(S_1 = 0, S_2 = 2, S_3 = 1) = \delta\mathbf{P}(\mathbf{0})\mathbf{\Gamma}\mathbf{P}(\mathbf{2})\mathbf{\Gamma}\mathbf{P}(\mathbf{1})\mathbf{1}'$$

$$= (4/13\ 9/13)\begin{pmatrix} 0.368 & 0 \\ 0 & 0.050 \end{pmatrix}\begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix}\begin{pmatrix} 0.184 & 0 \\ 0 & 0.0224 \end{pmatrix}\begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{pmatrix}\begin{pmatrix} 0.368 & 0 \\ 0 & 0.149 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \approx 0.007292$$

Exercicio 3

3. Considere um Modelo Oculto de Markov (HMM) $\{S_t, C_t : t \in \mathbb{N}\}$ de m-estados baseado em C_t, uma Cadeia de Markov estacionária com matriz de probabilidade de transição $\boldsymbol{\Gamma}$, distribuição estacionária $\boldsymbol{\delta} = (\delta_1, \delta_2, \cdots, \delta_m)$ e possuindo distribuições de estado dependentes univariadas $p_i(x)$. Sejam μ_i e σ_i^2 as médias e variâncias das distribuições p_i, seja ainda $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)$ o vetor de médias e $\mathbf{M} = \mathbf{diag}(\boldsymbol{\mu})$.

   Obtenha os seguintes resultados para os momentos do HMM.

   (a) $\mathrm{E}(S_t) = \sum_{i=1}^{m} \delta_i \mu_i = \boldsymbol{\delta}\boldsymbol{\mu}^\top$.

   (b) $\mathrm{E}(S_t^2) = \sum_{i=1}^{m} \delta_i (\sigma_i^2 + \mu_i^2)$.

Desenvolvimento A

Sabemos que se $X \sim P_0(\lambda)$ então $\mu = E(X) = Var(X) = \lambda$

$E(S_t) = \sum_{i=1}^m E(S_t \mid C_t = i) P(C_t = i) = \sum_{i=1}^m \lambda_i \delta_i = \delta\lambda'$

Desenvolvimento B

Sabemos que se $X \sim P_0(\lambda)$ então $E(X^2) = Var(X) + E(X)^2 = \lambda + \lambda^2$.

Partindo desse fato temos:

$E(S_t^2) = \sum_{i=1}^m E(S_t^2 \mid C_t = i) P(C_t = i)$

$= \sum_{i=1}^m (\lambda_i + \lambda_i^2)\delta_i = \sum_{i=1}^m \lambda_i^2 \delta_i + \sum_{i=1}^m \lambda_i \delta_i$

$= \lambda \mathbf{D} \lambda' + \delta\lambda'$

Exercicio 5

5. Considere $\{S_t, C_t : t \in \mathbb{N}\}$ um Modelo Oculto de Markov Poisson com três estados, com vetor de médias λ_i, $i = 1, 2, 3$ e matriz de probabilidade de transição

$$\Gamma = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 & 0 & 1/3 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

Assuma que a Cadeia de Markov é estacionária. Mostre que a função de autocorrelação $\rho(k) = \mathrm{Corr}(S_t, S_{t+k})$ é dada por

$$\rho(k) = \frac{(-1/3)^k \big(3(-5\lambda_1 - 3\lambda_2 + 8\lambda_3)^2\big) + k(-1/3)^{k-1}\big(4(-5\lambda_1 - 3\lambda_2 + 8\lambda_3)(\lambda_2 - \lambda_1)\big)}{32\big(15(\lambda_1^2 + \lambda_1) + 9(\lambda_2^2 + \lambda_2) + 8(\lambda_3^2 + \lambda_3)\big) - \big(15\lambda_1 + 9\lambda_2 + 8\lambda_3\big)^2}.$$

Note que, como uma função de k, esta expressão bastante tediosa é uma combinação linear de $(-1/3)^k$ e $k(-1/3)^{k-1}$.

Desenvolver da seguinte forma: simule um HMM Poisson com três estados, $\lambda = (1, 5, 10)$ e mesma matriz de probabilidades de transição apresentada. Estime a função de autocorrelação da Cadeia de MArkov subjacente e interprete.

Desenvolvimento

Modelo gerado pela função *hmmspec* do pacote *mhsmm* com valores de Γ e $\lambda's$ solicitados:

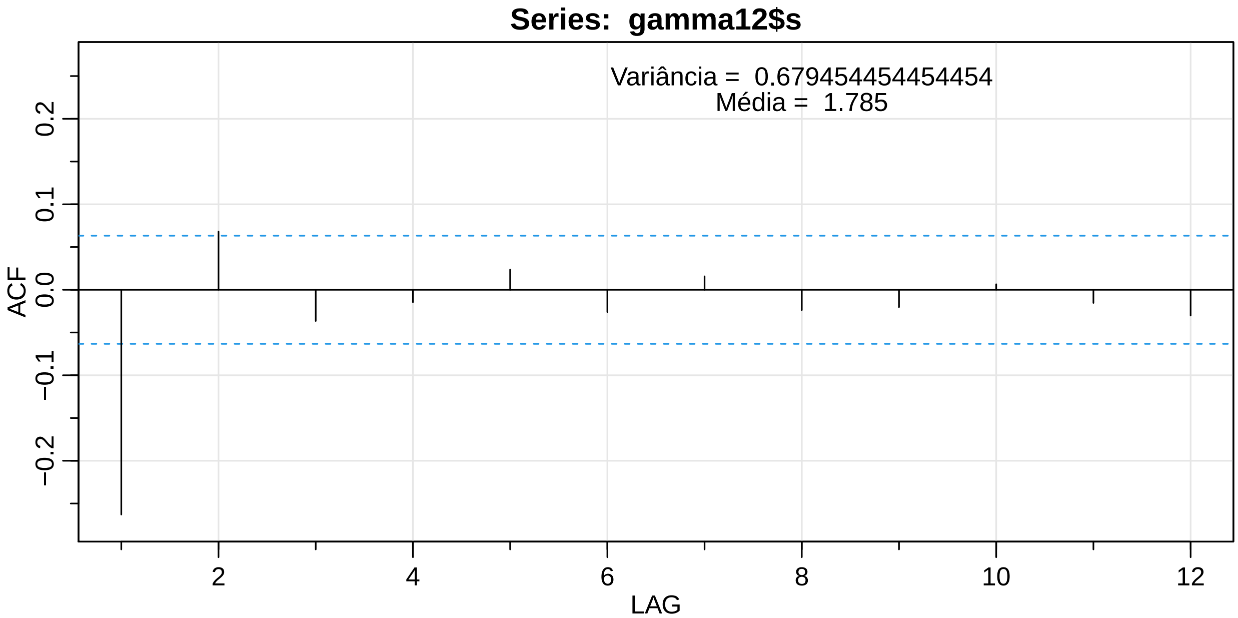
```
Hidden Markov Model specification:
J (number of states):
3
init:
[1] 0.3333 0.3333 0.3333
transition:
       [,1]   [,2]   [,3]
[1,] 0.3333 0.3333 0.3333
[2,] 0.6667 0.0000 0.3333
[3,] 0.5000 0.5000 0.0000
emission:
$lambda
[1]  1  5 10
```

$\lambda's$ estimados via simulação com 1000 repetições:

```
[1] 1.045
[1] 5.044
[1] 9.627
```

Plotando a função de autocorrelação da série simulada:

```
 [1] -0.26  0.07 -0.04 -0.01  0.02 -0.03  0.02 -0.02 -0.02  0.01 -0.02 -0.03
```



Sabemos que se $X \sim P(\lambda_i)$ então $Var(X) = E(X) = \lambda$. Os valores estimados da média e variância não são próximos, indicando que a cadeia pode ser uma mistura de séries. E como foi dado no enunciado, a distribuição é Poisson, ou seja, a cadeia foi obtida como mistura de três variáveis Poissons.