Lista 2 - MAC 6926/MAE 0580

Setembro de 2017

Notação e desigualdade de Chebyshev

Dado $n \geq 1$ e dada uma amostra $X_{-k}, \cdots, X_n$, definimos a função de contagem

$$N_n(a_{-k}^0) = \sum_{t=0}^{n} I_{\{X_{t-k}^t = a_{-k}^0\}}$$

para toda sequência $a_{-k}^0 \in A^{k+1}$.

Desigualdade de Chebyshev. Para uma variável aleatória Z com média $\mathbb{E}(Z)$ e variância $\mathrm{Var}(Z)$ finitas, vale que

$$\mathbb{P}(|Z - \mathbb{E}(Z)| > \epsilon) \leq \frac{\mathrm{Var}(Z)}{\epsilon^2}$$

para todo $\epsilon > 0$.

1. Dada a amostra

   $$X_{-1} = 0, X_0 = 1, X_1 = 0, X_2 = 0, X_3 = 0, X_4 = 1, X_5 = 0, X_6 = 1, X_7 = 1, X_8 = 1, X_9 = 0, X_{10}$$

   (a) calcule a verossimilhança, supondo que $P(X_{-1} = 0) = 1$ e que $(X_t)_{t \geq -1}$ é uma cadeia de Markov de ordem 1 com matriz de transição P dada por:

   $$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$

   (b) Suponha agora que $(X_t)_{t \geq -1}$ é uma cadeia de Markov de alcance 2 com matriz de probabilidade de transição p dada por:

a	$p(a\|00)$	$p(a\|01)$	$p(a\|10)$	$p(a\|11)$
0	0.7	0.5	0.4	0.8
1	0.3	0.5	0.6	0.2

Calcule a verossimilhança da amostra dado que $X_{-1} = 0$ e $X_0 = 1$.

(c) Determine $\hat{p}$ que maximiza a verossimilhança para ambos os casos acima.

2. Considere uma realização aleatória $(X_n)_{n=0,\ldots,100}$ de uma cadeia de Markov assumindo valores no alfabeto $A = \{0,1\}$ e com matriz de probabilidades de transição

$$p = \begin{pmatrix} 1/4 & 3/4 \\ 3/5 & 2/5 \end{pmatrix} .$$

Obteve-se as contagens de todas as sequências de tamanho 2 na amostra:

$$N_{100}(00) = 15, N_{100}(01) = 48, N_{100}(10) = 21, N_{100}(11) = 16 .$$

Assumindo que $X_0 = 1$ é dado:

(a) Calcule a verossimilhança da amostra.

(b) Obtenha os estimadores de máxima verossimilhança das probabilidades de transição da matriz p.

(c) Calcule o *maior valor* que a verossimilhança da amostra pode assumir.

3. Seja $(Z_n)_{n\geq 0}$ uma cadeia de Markov de alcance 1 assumindo valores no alfabeto $B = \{0,1\}^2$ e com matriz de probabilidades de transição Q dada por

$$\begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{pmatrix} 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 1/6 & 5/6 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 3/4 & 1/4 \end{pmatrix} \end{array} .$$

Nessas condições, verifique que a matriz Q define uma cadeia de alcance variável que assume valores no alfabeto $A = \{0,1\}$ e com árvore de contextos dada por

$$\tau = \{\{w_{-1} = 0\}, \{w_{-2} = 0, w_{-1} = 1\}, \{w_{-2} = 1, w_{-1} = 1\}\}.$$

4. Sejam $k, n \geq 1$ dois números inteiros, $X_{-k}^{n} = (X_{-k}, \ldots, X_n)$ uma amostra de uma cadeia com memória de alcance variável definida em um alfabeto A finito com árvore de contextos τ. Suponhamos que $X_{-k}^{-1} = w \in \tau$, isto é, X_{-k}^{-1} é um contexto da árvore τ, definimos para cada $a \in A$,

$$N_n(wa) = \sum_{t=0}^{n} \mathbf{1}\{X_{t-|w|}^{t-1} = w, X_t = a\},$$

como sendo o número de vezes que o símbolo a é precedido pelo contexto w na amostra X_0^n.

Seja (X_k) uma cadeia com memória de alcance variável, assumindo valores no alfabeto finito $A = \{0, 1\}$, tendo como árvore de contextos

$$\tau = \{\{w_{-1} = 1\}, \{w_{-2} = 1, w_{-1} = 0\}, \{w_{-2} = 0, w_{-1} = 0\}\}$$

e tendo família associada de probabilidades de transição definida por

$$p(1 \,|\, 1) = \alpha \,, p(1 \,|\, 01) = \beta \;\text{ e }\; p(1 \,|\, 00) = \gamma \,,$$

onde α, β e γ são tres parâmetros pertencentes ao intervalo aberto $(0, 1)$.

(a) Suponhamos que a amostra gerada foi

$$X_{-1}^{10} = x_{-1}^{10} = (1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0).$$

Supondo que $\mathbb{P}(X_{-1} = 1) = 1$, diga quanto vale $\mathbb{P}(X_1^{10} = x_{-1}^{10})$.

(b) Verifique que $\sum_{a \in A} N_n(wa) = N_{n-1}(w)$.

(c) Verifique que $\sum_{w \in \tau} \sum_{a \in A} N_n(wa) = n + 1$

5. Nas condições do Exercício 5, encontre as estimativas de máxima verossimilhança dos parâmetros α, β e γ.

6. Dada a amostra

$$X_{-1} = 0, X_0 = 1, X_1 = 0, X_2 = 0, X_3 = 0, X_4 = 1, X_5 = 0,$$
$$X_6 = 1, X_7 = 1, X_8 = 1, X_9 = 0, X_{10} = 1.$$

(a) Determine $\hat{p} \in \mathcal{M}_0(\{0, 1\})$ que maximiza a verossimilhança da amostra.

(b) Calcule a verossimilhança da amostra, supondo que $X_{-1} = 0$ dado e que $(X_t)_{t \geq -1}$ é uma cadeia de Markov de ordem 1 assumindo valores no alfabeto $A = \{0, 1\}$, com matriz de transição P dado por

$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$

Em seguida, determine $\hat{p} \in \mathcal{M}_1(A)$ que maximiza a verossimilhança da amostra.

(c) Suponha agora que $X_{-1} = 0$ e $X_0 = 1$ são dados e que $(X_t)_{t \geq -1}$ é uma cadeia de Markov de alcance 2 assumindo valores no alfabeto $A^2 = \{0, 1\}^2$, com matriz de transição $p = \{p(a|x_{-2}^{-1}) : a \in A, x_{-2}^{-1} \in A^2\}$ com

a	$p(a\|00)$	$p(a\|01)$	$p(a\|10)$	$p(a\|11)$
0	0.7	0.5	0.4	0.8
1	0.3	0.5	0.6	0.2

Calcule a verossimilhança da amostra. Em seguida, determine $\hat{p} \in \mathcal{M}_2(A^2)$ que maximiza a verossimilhança.

7. Seja $(X_n)_{n \geq -2}$ uma evolução Markoviana com memória de alcance 2 assumindo valores no alfabeto $A = \{0, 1\}$ e que pode ser simulada através do seguinte algoritmo:

Passo 1. $X_{-2} = 1$ e $X_{-1} = 0$;

Passo 2. Para $n \geq 0$, definimos

$$X_n = \begin{cases} 0, & \text{se } U_n \leq h(X_{n-2}, X_{n-1}) \\ 1, & \text{se } U_n > h(X_{n-2}, X_{n-1}) \end{cases}$$

onde $h(0,0) = 1/2$, $h(0,1) = 1/3$, $h(1,0) = 1/4$ $h(1,1) = 1/5$ e $(U_n)_{n \geq 0}$ é uma sequência de variáveis aleatórias i.i.d com distribuição uniforme em $[0, 1]$.

(a) Qual é a matriz de probabilidades de transição desta cadeia de Markov de alcance 2?

(b) Calcule $\mathbb{P}(X_1 = 1)$.

8. Considere a cadeia estocástica $(X_n)_{n\geq -2}$ definida no exercício anterior. Seja $(Y_n)_{n\geq 0}$ a cadeia estocástica tomando valores no alfabeto $S = \{0,1\}^2$ satisfazendo $Y_n = (X_{n-2}, X_{n-1})$.

   (a) Observe que $(Y_n)_{n\geq 0}$ é uma cadeia de Markov de ordem 1.

   (b) Determine a matriz de transição desta cadeia de Markov.

   (c) O que podemos dizer a respeito de cadeias de Markov de alcance 1 em A^k construidas a partir de cadeias de alcance k em A?

9. Seja $(X_n)_{n\geq 1}$ uma sequência de variáveis aleatórias iid, assumindo valores no alfabeto $A = \{0,1\}$, tal que $\mathbb{P}(X_n = 1) = p$, onde $0 < p < 1$. Usando a desigualdade de Chebyshev, qual o menor $\overline{n} \geq 1$ tal que, para todo $n \geq \overline{n}$, tenhamos:

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - p\right| > 0.01\right) \leq 0.01$$

10. Seja $(X_n)_{n\geq 0}$ uma sequência de símbolos pertencentes ao alfabeto finito $A = \{0,1\}$ e gerada por uma cadeia com memória de alcance variável com árvore de contextos

$$\tau = \{\{X_{-1} = 0\}, \{X_{-2} = 0, X_{-1} = 1\}, \{X_{-2} = 1, X_{-1} = 1\}\}$$

   e família de probabilidades de transição p desconhecida. Dada a amostra

$$X_{-1}^{22} = (0,0,1,1,1,0,1,0,1,0,1,1,0,1,1,0,0,1,0,1,1,1,0,1)\,.$$

   gerada por esta cadeia

   a) calcule a verossimilhança da amostra.

   b) Estime as probabilidades de transi c cão pelo método da máxima verossimilhan c ca.

11. Seja $X_1^{1000} = (X_1, \cdots, X_{1000})$ uma realização de uma cadeia de Markov de alcance 2 assumindo valores no alfabeto $A = \{0,1\}$. A partir da amostra, obteve-se os valores das seguintes funções de contagem:

$$N_{1000}(000) = 150;\ N_{1000}(001) = 54;\ N_{1000}(010) = 167;\ N_{1000}(011) = 116;$$
$$N_{1000}(100) = 55;\ N_{1000}(101) = 229;\ N_{1000}(110) = 116.$$

Estime a matriz de probabilidade $\hat{p}$ pelo método da máxima verossimilhança.

12. Seja $(X_k)_{k\geq 0}$ uma cadeia com memória de alcance variável, assumindo valores no alfabeto $A = \{0,1\}$, tendo como árvore de contextos $\tau = \{\{X_{-1}=0\}, \{X_{-2}=0, X_{-1}=1\}, \{X_{-2}=1, X_{-1}=1\}\}$ e tendo família associada de probabilidades de transição definida por

$$P(X_n = 1 \,|\, X_{n-1} = 0) = \alpha$$
$$P(X_n = 1 \,|\, X_{n-1} = 1, X_{n-2} = 0) = \beta$$
$$P(X_n = 1 \,|\, X_{n-1} = 1, X_{n-2} = 1) = \gamma\,,$$

onde α, β e γ são tres parâmetros pertencentes ao intervalo aberto $(0,1)$. A partir de uma amostra $X_{-1}^{12} = (0,0,1,1,1,0,0,0,1,0,1,1,1,1)$, gerada por esta cadeia e supondo que $P(X_{-1}=0)=1$,

a) Calcule a verossimilhança da amostra.

b) Estime os parâmetros $\hat{\alpha}, \hat{\beta}$ e $\hat{\gamma}$ pelo método da máxima verossimilhança.

13. Seja $(X_0, \ldots, X_n)$ uma amostra de uma cadeia de alcance variável assumindo valores no alfabeto $A = \{0,1\}$, com árvore de contextos $\tau = \{\{X_{-1}=0\}, \{X_{-2}=0, X_{-1}=1\}, \{X_{-2}=1, X_{-1}=1\}\}$ e com família de probabilidades de transição p dada por

$$p(0\,|\,0) = 0,4\,, p(0\,|\,10) = 0,2 \text{ e } p(0\,|\,11) = 0,6\,.$$

Nessas condições, quanto vale

$$\lim_{n\to\infty} \frac{N_n(000)}{N_{n-1}(00)}?$$