MAC 6926 - MAE 0580 - Prova 2 (Turma A)

18 de outubro de 2017

Esta é uma prova individual, sem consulta. Em cada uma das 10 questões abaixo, uma e só uma opção é correta. A nota da prova ser calculada pela fórmula Nota = $\max\{0, C - E/3\}$. Nesta expressão, C é o número de respostas certas e E, o número de respostas erradas. Questões deixadas em branco e respostas rasuradas *não serão consideradas* no cálculo da nota.

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

$$N_{0:n}(a_{-k}^0) = \sum_{t=0}^{n} \mathbb{1}_{\{X_{t-k}^t = a_{-k}^0\}}$$

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

Dada uma amostra $X_{-k}, \cdots, X_n$ de símbolos no alfabeto A com árvore de contextos τ com altura k, definimos

$$\hat{\mathbb{P}}_\tau(X_0^n | X_{-k}^{-1}) = \prod_{\omega \in \tau} \prod_{a \in A} \hat{p}_n(a|\omega)^{N_{0:n}(\omega a)},$$

onde $\hat{p}_n(a|\omega) = \frac{N_{0:n}(\omega a)}{N_{0:n-1}(\omega)}$ é o estimador de máxima verossimilhança da matriz de probabilidades de transição, dada a árvore τ.

Desigualdade de Hoeffding para variáveis aleatórias binárias i.i.d.: Sejam $Y_1, Y_2, \cdots$ assumindo valores em {0,1} com $\mathbb{P}(Y_n = 1) = p$. Então para todo $\delta > 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i > n(p + \delta)\right) \leq exp\{-2n\delta^2\}$$

Algoritmo Contexto:
Dada uma amostra $X_{-k}^n = (X_{-k}, \ldots, X_n)$ de símbolos do alfabeto finito, definimos para toda sequência $w = w_{-k}^{-1} \in A^k$, com $1 < k < n$, a seguinte quantidade

$$\Delta_n(w_{-k}^{-1}) = \max_{a,b \in A} \left| \hat{p}_n(a|w_{-k}^{-1}) - \hat{p}_n(a|w_{-k}^{-1}b) \right| .$$

Fixado $\delta \in (0, 1)$, se $\Delta_n(w_{-k}^{-1}) < \delta$, então podamos os símbolos mais remotos (representados pela letra b) das sequências $\{bw_{-k}^{-1} : b \in A\}$. Caso contrário, mantemos as sequências $\{bw_{-k}^{-1} : b \in A\}$.

Sejam (X, Y) duas variáveis aleatórias com $X \in \mathcal{X}$ e $Y \in \{0, 1\}$ e seja $\mathcal{F}$ uma classe de funções de $\mathcal{X}$ em $\{0, 1\}$. Dada $f \in \mathcal{F}$, definimos $R(f) = \mathbb{P}(f(X) \neq Y)$.
Dada $(X_i, Y_i), i = 1, \cdots, n$ uma sequência de variáveis aleatórias i.i.d. com a mesma distribuição de (X, Y), definimos o risco empírico

$$R_n(f) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{f(X_i) \neq Y_i\}}.$$

Para toda sequência $x_1, \cdots, x_n$ definimos

$$\mathcal{N}_{\mathcal{F}}(x_1, \cdots, x_n) = \{f(x_1), \cdots, f(x_n) : f \in \mathcal{F}\}.$$

Definimos também o coeficiente de fragmentação

$$\mathcal{S}(\mathcal{F}, n) = \max_{(x_1,\cdots,x_n)\in\mathcal{X}^n} |\mathcal{N}_{\mathcal{F}}(x_1, \cdots, x_n)|.$$

A dimensão de Vapnik-Chervonenkis da classe $\mathcal{F}$ é definida como

$$VC(\mathcal{F}) = \max\{n \geq 1 : \mathcal{S}(\mathcal{F}, n) = 2^n\}.$$

Desigualdades VC:

$$\mathbb{P}(\sup_{f\in\mathcal{F}} |R_n(f) - R(f)| > \epsilon) \leq 8\mathcal{S}(\mathcal{F}, n)e^{-n\epsilon^2/32}$$

e

$$\mathbb{E}\left(\sup_{f\in\mathcal{F}} |R_n(f) - R(f)|\right) \leq 2\sqrt{\frac{\log \mathcal{S}(\mathcal{F}, n) + \log 2}{n}}$$

Lema de Sauer:

$$\mathcal{S}(\mathcal{F}, n) \leq (n+1)^{VC(\mathcal{F})}$$

1. Queremos simular a cadeia com memória de alcance variável $(X_n)_{\geq -2}$, 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

$$\begin{aligned}\mathbb{P}(X_n = 0 \,|\, X_{n-1} = 0) &= 0,5\\ \mathbb{P}(X_n = 0 \,|\, X_{n-1} = 1, X_{n-2} = 0) &= 0,7\\ \mathbb{P}(X_n = 0 \,|\, X_{n-1} = 1, X_{n-2} = 1) &= 0,2.\end{aligned}$$

   Fazemos isso com o seguinte algoritmo de simulação:

   1. Inicialização: $X_{-2} = X_{-1} = 1$.

   2. Para todo $n \geq 0$, $X_n = \mathbb{1}_{\{U_n > p(0|c_\tau(X_{-2}^{n-1}))\}}$, onde $c_\tau(X_{-2}^{n-1})$ o contexto associado por τ à sequência X_{-2}^{n-1}.

   Sorteando $U_0 = 0,15$, $U_1 = 0,44$, $U_2 = 0,83$ e $U_3 = 0,27$, diga qual das sequências abaixo foi produzida pelo algoritmo.

   (a) $X_0 = 1$, $X_1 = 1$, $X_2 = 0$, $X_3 = 0$.
   (b) $X_0 = 0$, $X_1 = 0$, $X_2 = 1$, $X_3 = 0$.
   (c) $X_0 = 0$, $X_1 = 0$, $X_2 = 1$, $X_3 = 1$.
   (d) Nenhuma das respostas anteriores.

2. 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 Hoeffding, diga qual das afirmações é verdadeira (use se necessário $\log(0.02) = -3.91$)

   (a) $\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - p > 0.05\right) \leq 0.02, \forall n \geq 300$

   (b) $\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - p > 0.02\right) \leq 0.02, \forall n \geq 4.000$

   (c) $\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - p > 0.01\right) \leq 0.02, \forall n \geq 20.000$

   (d) Nenhuma das respostas anteriores.

3. Seja $X_{-2}^{1000} = (X_{-2}, X_{-1} \cdots, X_{1000})$ uma realizaão 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 funes de contagem:

$$N_{0:1000}(001) = 54;\ N_{0:1000}(010) = 167; N_{0:1000}(011) = 116;$$
$$N_{0:1000}(100) = 55;\ N_{0:1000}(101) = 229;\ N_{0:1000}(110) = 116;\ N_{0:1000}(111) = 150.$$

   O valor correto da probabilidade de transião estimada por máxima verossimilhança :

   a) $\hat{p}(0\,|\,00) = 114/167$

   b) $\hat{p}(0\,|\,00) = 113/168$

   c) $\hat{p}(0\,|\,00) = 114/168$

   d) Nenhuma das respostas anteriores.

4. Dada a amostra $(0, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1)$, diga qual é a árvore de contextos $\hat{\tau}$ de altura menor ou igual a 2, obtida a partir da aplicaão do Algoritmo Contexto, utilizando $\delta = 0.05$ no critério de poda.

   (a) $\hat{\tau} = \{1, 0\}$

   (b) $\hat{\tau} = \{1, 00, 10\}$

   (c) $\hat{\tau} = \{0, 11, 01\}$

   (d) Nenhuma das respostas anteriores.

5. Seja $\mathcal{X} = [0, 1]$ e seja $\mathcal{F}$ o conjunto de funções de $\mathcal{X}$ em $\{0, 1\}$ assim definida: $f \in \mathcal{F}$ se

$$f(x) = \mathbf{1}\{a_1 \leq x \leq b_1\} + \mathbf{1}\{a_2 \leq x \leq b_2\}$$

   com $0 \leq a_1 < b_1 < a_2 < b_2 \leq 1$. Diga qual das afirmações abaixo é verdadeira:

   (a) $\mathcal{S}(\mathcal{F}, 5) < 2^5$

   (b) Para toda sequência $((x_i, y_i) : i = 1, \cdots, 6)$ com $x_i \in \mathcal{X}$ e $y_i \in \{0, 1\}$ existe $f \in \mathcal{F}$ tal que $y_i = f(x_i), i = 1, \cdots, 6$.

   (c) $VC(\mathcal{F}) = 5$

   (d) Nenhuma das anteriores.

6. Seja $\mathcal{X} = [0,1]$ e seja $\mathcal{F}$ o conjunto de funções de $\mathcal{X}$ em $\{0,1\}$ assim definida: $f \in \mathcal{F}$ se

$$f(x) = \mathbf{1}\{x < t\} \text{ ou } f(x) = \mathbf{1}\{x \geq t\}$$

para algum $t \in [0,1]$. Diga qual das seguintes opções é verdadeira para todo $n \geq 2$

(a) $\mathbb{E}\left(\sup_{f\in\mathcal{F}} |R_n(f) - R(f)|\right) \leq 2\sqrt{\frac{\log 2(n+1)^2}{n}}$

(b) $\mathbb{E}\left(\sup_{f\in\mathcal{F}} |R_n(f) - R(f)|\right) \leq 2\sqrt{\frac{\log 4(n+1)}{n}}$

(c) $\mathbb{E}\left(\sup_{f\in\mathcal{F}} |R_n(f) - R(f)|\right) \leq 2\sqrt{\frac{\log (n+1)^2}{n}}$

(d) Nenhuma das anteriores.

7. Usando a notação introduzida no preâmbulo, diga qual das seguintes opções é verdadeira para todo $\delta \in (0, 1/2)$

(a) $\sup_{f\in\mathcal{F}} |R_n(f) - R(f)| < 4\sqrt{\frac{2}{n}\log\frac{8\mathcal{S}(\mathcal{F},n)}{\delta}}$, com probabilidade menor ou igual a δ.

(b) $\sup_{f\in\mathcal{F}} |R_n(f) - R(f)| < 4\sqrt{\frac{2}{n}\log\frac{8\mathcal{S}(\mathcal{F},n)}{\delta}}$, com probabilidade maior ou igual a $1-\delta$.

(c) $\sup_{f\in\mathcal{F}} |R_n(f) - R(f)| < 4\sqrt{\frac{2}{n}\log\frac{8\mathcal{S}(\mathcal{F},n)}{\delta}}$, com probabilidade igual a 8δ.

(d) Nenhuma das anteriores.

8. Seja $\xi_1, \xi_2, \cdots, \xi_n$ variáveis aleatórias i.i.d. assumindo valores em $\{-1, 1\}$ com $\mathbb{P}(\xi_n = -1) = \frac{1}{2}$. Usando a desigualdade de Hoeffding, diga qual das afirmativas abaixo é verdadeira

(a) $\mathbb{P}(\sum_{i=1}^n \xi_i > n\epsilon) \leq \exp\{-n\epsilon^2/2\}$.

(b) $\mathbb{P}(\sum_{i=1}^n \xi_i > n\epsilon) \leq \exp\{-n\epsilon^2\}$.

(c) $\mathbb{P}(\sum_{i=1}^n \xi_i > n\epsilon) \leq \exp\{-2n\epsilon^2\}$.

(d) Nenhuma das anteriores.