Notas práctico 6 parte 2

viernes, 11 de junio de 2021 14:03

Ejercicio 3

$$-\log Q_{NML}(x^n) + \log P_{ML}(x^n) = \text{cte}$$

$$\Rightarrow \max_{x^n \in \mathcal{X}^n} \left\{ -\log Q_{NML}(x^n) + \log P_{ML}(x^n) \right\} = \text{cte.}$$

Sea Q distr. sobre $\mathcal{X}^n$, $Q \neq Q_{NML}$

Existe $y^n \in \mathcal{X}^n$ t.q. $Q(y^n) < Q_{NML}(y^n)$

$$\Rightarrow -\log Q(y^n) + \log P_{ML}(y^n) > -\log Q_{NML}(y^n) + \log P_{ML}(y^n)$$

$$= \max_{x^n \in \mathcal{X}^n} \left\{ -\log Q_{NML}(x^n) + P_{ML}(x^n) \right\}$$

Ejercicio 4.1

Códigos universales con tasa de convergencia óptima

$$-\log Q(x^n) = -\log P_{ML}(x^n) + \underbrace{\frac{|S|(|\mathcal{X}|-1)}{2} \log n + \Theta(1)}_{\text{Reg.}}$$

$$-\frac{\log Q(x^n)}{n} = \underbrace{\hat{H}(x^n)}_{\substack{\text{término}\\ \text{de ajuste}}} + \underbrace{\frac{|S|(|\mathcal{X}|-1)}{2} \frac{\log n}{n} + \Theta(1/n)}_{\substack{\text{costo de modelo}\\ \text{más riqueza} \Rightarrow \text{más grande}}}$$

Más riqueza $\Rightarrow$ más chico

Ejemplo

$$x^n = 0101010101 \ldots\ldots 01$$

* Con $k = 0$, $\hat{H}_k(x^n) = 1$ bit/símbolo

  costo modelo $= \frac{1}{2} \frac{\log n}{n} + \Theta(1/n)$

* Con $k = 1$, $\hat{H}_k(x^n) = 0$ bit/símbolo

  costo modelo $= 1 \cdot \frac{\log n}{n} + \Theta(1/n)$

* Con $k = 2$, $\hat{H}_k(x^n) = 0$ bit/símbolo

  costo modelo $= 2 \frac{\log n}{n} + \Theta(1)$

---

asignación KT de orden 0.

$$Q_n^{KT,k}(x^n) = \prod_{s \in S} Q_{n_s}^{KT}(z_s),$$

asignación KT para modelos Markovianos de orden k

donde z_s es la secuencia de símbolos de x^n que ocurre en estado s, y n_s es el largo de z_s

$$-\log Q_n^{KT,k}(x^n) = \sum_{s \in S} \underbrace{-\log Q_{n_s}^{KT}(z_s)}$$

$$\leq -\log \hat{P}(z_s) + \frac{1}{2} \log n_s + \Theta(1)$$

de orden 0.

$$\leq -\log \hat{P}(z_s) + \frac{1}{2} \log n$$

$$\leq |S| \cdot \left(\frac{1}{2} \log n + \Theta(1) \right) + \underbrace{\sum_{s \in S} -\log \hat{P}(z_s)}_{-\log P_{ML}(x^n)}$$

Markov de orden k

$$P_{ML}(x^n) = \prod_{s \in S} \hat{P}(z_s)$$

$\Theta = (\theta_s)_{s \in S}$, donde θ_s es la prob. de 0 en estado s

vector de tamaño $|S|$.

$$P_\Theta(x^n) = \prod_{s \in S} \theta_s^{n_{0|s}} (1-\theta_s)^{n_{1|s}}$$

Se maximiza haciendo que θ_s maximice $\theta_s^{n_{0|s}} (1-\theta_s)^{n_{1|s}}$, para cada $s \in S$

Y este factor es máximo cuando

$$\theta_s = \frac{n_{0|s}}{n_{0|s} + n_{1|s}}$$