模式识别 第二次作业

2021 年 10 月 7 日

说明

- 作业用中文撰写，务必注明题号，鼓励使用 LaTex。
- 文档按“学号 _ 姓名.pdf”命名提交。
- 本次作业截止时间为 2021 年 10 月 20 日，请到课程网站及时提交。

1. 设 x 的概率密度为均匀分布：

$$p(x|\theta) \sim U(0,\theta) = \begin{cases} 1/\theta & 0 \leq x \leq \theta \\ 0 & otherwise \end{cases}$$

(1) 假设 n 个样本 $\mathcal{D} = \{x_1, ..., x_n\}$ 都独立地服从分布 $p(x|\theta)$。证明对于 θ 的最大似然估计就是 $\mathcal{D}$ 中的最大值 $\max\ [\mathcal{D}]$。

(2) 假设从该分布中采样 5 个样本 ($n = 5$)，且有 $\max\limits_k x_k = 0.6$, 画出在区间 $0 \leq \theta \leq 1$ 上的似然函数 $p(\mathcal{D}|\theta)$, 并解释为什么此时不需要知道其余四个点的值。

2. 设一组样本服从高斯分布，其协方差矩阵 $\boldsymbol{\Sigma}$ 已知，但均值 $\boldsymbol{\mu}$ 未知。假设该均值 $\boldsymbol{\mu}$ 的先验分布为 $\mathcal{N}(\boldsymbol{m}_0, \boldsymbol{\Sigma}_0)$。

(1) 均值 $\boldsymbol{\mu}$ 的 MAP 估计是什么。

(2) 若使用线性变换对样本进行变换：$\boldsymbol{x}^{'} = \boldsymbol{A}\boldsymbol{x}$, 其中 $\boldsymbol{A}$ 为非奇异矩阵。请问 MAP 能否正确地对变换之后的 $\boldsymbol{\mu}^{'}$ 进行估计？为什么？

3. 数据 $\mathcal{D} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ * \end{pmatrix} \right\}$ 中的样本独立地服从二维的分布 $p(x_1, x_2) = p(x_1)p(x_2)$。其中 $*$ 代表丢失的数据，且有

$$p(x_1) \sim \begin{cases} \dfrac{1}{\theta_1} e^{-x_1\theta_1} & x_1 \geq 0 \\ 0 & otherwise \end{cases}$$

和

$$p(x_2) \sim U(0, \theta_2) = \begin{cases} \dfrac{1}{\theta_2} & 0 \leq x_2 \leq \theta \\ 0 & otherwise \end{cases}$$

(1) 假设初始估计为 $\boldsymbol{\theta}^0 = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$，计算 $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^0)$(EM 算法中的 E 步)。注意要对分布进行归一化。

(2) 求使得 $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^0)$ 最大的那个 $\boldsymbol{\theta}$(EM 算法中的 M 步)。

4. 用前向-后向算法训练一个 HMM，已知训练序列长为 T，其中每一时刻都可能取 c 个符号中的一个。那么全部更新一次 $\hat{a}_{ij}$ 和 $\hat{b}_{jk}$ 的计算复杂度是多少?

5. 假设有正态分布 $p(x) \sim \mathcal{N}(\mu, \sigma^2)$ 和 Parzen 函数 $\varphi(x) \sim \mathcal{N}(0, 1)$。证明 Parzen 窗估计

$$p_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \varphi\left(\frac{x - x_i}{h_n}\right)$$

有如下性质:

(1) $\bar{p}_n(x) \sim \mathcal{N}(\mu, \sigma^2 + h_n^2)$。

(2) $\mathrm{Var}[p_n(x)] \approx \frac{1}{2nh_n\sqrt{\pi}} p(x)$。