Machine Learning and Computational Statistics Theodoros Georgiopoulos

Homework 5

Exercises for Unit 5: Pdf estimation - Inference

Solution for exercise 1

We have the Erlang distribution $p(x) = \theta^2 x \exp(-\theta x) u(x)$, where u(x) = 1(0), if $x \ge 0 (< 0)$, with $E[x] = 2/\theta$.

(a) Given a set of N measurements we can calculate the ML estimate of θ as follows. Firstly, we calculate the log-likelihood

$$L(\theta) = \ln \left(\prod_{i=1}^{N} \theta^2 x_i \exp(-\theta x_i) u(x_i) \right)$$
$$= \sum_{i=1}^{N} \ln \left(\theta^2 x_i \exp(-\theta x_i) u(x_i) \right)$$

For $x \geq 0$,

$$L(\theta) = \sum_{i=1}^{N} \ln \theta^{2} + \sum_{i=1}^{N} \ln x_{i} - \theta \sum_{i=1}^{N} x_{i}$$
$$= 2N \ln \theta + \sum_{i=1}^{N} \ln x_{i} - \theta \sum_{i=1}^{N} x_{i}.$$

We find the derivative of the log-likelihood with respect to θ

$$\frac{\partial L(\theta)}{\partial \theta} = \frac{2N}{\theta_{ML}} - \sum_{i=1}^{N} x_i$$

and we set it to zero. So,

$$\frac{\partial L(\theta)}{\partial \theta} = 0 \iff \theta_{ML} = \frac{2N}{\sum_{i=1}^{N} x_i}.$$

(b) For N=5 we have $x_1=2, x_2=2.2, x_3=2.7, x_4=2.4, x_5=2, 6$. From (a), we have

$$\theta_{ML} = \frac{10}{11.9} = 0.84$$

and

$$E[x] = \frac{2}{\theta_{ML}} = \frac{2}{0.84} = 2.38$$

Solution for exercise 2

We have again the Erlang distribution $p(x) = \theta^2 x \exp(-\theta x) u(x)$, where u(x) = 1(0), if $x \geq 0 (< 0)$, a set of N measurements x_1, \ldots, x_n for the random variable x that follows Erlang distribution, and we know the a priori probability for θ , such that $\theta \sim \mathcal{N}(\theta_0, \sigma_0^2)$ with θ_0 and σ_0 known.

(a) The MAP estimate is

$$\hat{\theta}_{MAP} = \arg\max_{\theta} p(\theta) p(x|\theta).$$

It is convenient to maximize the logarithm of the product. So,

$$\begin{split} \hat{\theta}_{MAP} &= \arg \max_{\theta} \ln \left(p(\theta) p(x|\theta) \right) \\ &= \arg \max_{\theta} \left(\ln p(\theta) + \ln p(x|\theta) \right) \\ &= \arg \max_{\theta} \left(\ln p(\theta) + \ln \sum_{i=1}^{N} p(x_i|\theta) \right). \end{split}$$

We find the derivative of the sum and we set it to zero. For the second term, we have calculated it for the exercise 1.

$$-\frac{1}{2\sigma_0^2} 2(\theta_{MAP} - \theta_0)(-1) + \frac{2N}{\theta_{MAP}} - \sum_{i=1}^N x_i = 0$$

$$\frac{\theta_{MAP} - \theta_0}{\sigma_0^2} + \frac{2N}{\theta_{MAP}} - \sum_{i=1}^N x_i = 0$$

$$\theta_{MAP}^2 - \theta_0 \theta_{MAP} + 2N\sigma_0^2 - \theta_{MAP} \sum_{i=1}^N x_i = 0$$
(1)

Finally,

$$\theta_{MAP} = \frac{1}{2} \left((\theta_0 - \sigma_0^2 \sum_{i=1}^N x_i) + \sqrt{(\theta_0 - \sigma_0^2 \sum_{i=1}^N x_i)^2 + 8N\sigma_0^2} \right).$$

(b) For $N \to \infty$, if we divide (1) by N we observe that

$$\theta_{MAP} \to \frac{2N}{\sum_{i=1}^{N} x_i} = \theta_{ML}.$$

For $\sigma_0^2 >>$, from (1) we observe the same as before

$$\theta_{MAP} \to \frac{2N}{\sum_{i=1}^{N} x_i} = \theta_{ML}.$$

For $\sigma_0^2 \ll$, if we multiply (1) by σ_0^2 , we observe that

$$\theta_{MAP} \to \theta_0$$
.