COMS20011: Symbols, Patterns and Signals

Problem Sheet: Maximum likelihood

1. You were consulted by a Physics student who is trying to estimate the voltage (V) given current (I) and resistance (R) information. The student informs you that the physical model is
$$I = \frac{V}{R} + \epsilon$$
subject to a measurement error $\epsilon \sim \mathcal{N}(0, \sigma^2)$. Assuming i.i.d observations, the likelihood of $p(D|V)$ where V is the model's only parameter and D is the observed data is equal to:
$$p(D|V) = \prod_i \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(I_i - \frac{V}{R_i})^2}{\sigma^2}}$$
where I_i is the current value for observation i and R_i is the resistance value for observation i. Prove, using the Maximum Likelihood Estimation (MLE) recipe, that
$V_{ML} = \sum \frac{I_i}{R_i} / \sum \frac{1}{R_i^2}$
Answer:
1. Take the natural logarithm: $\ln p(D|V) = N \ln \frac{1}{\sqrt{2\pi}\sigma} + \sum_i -\frac{1}{2}\frac{(I_i - \frac{V}{R_i})^2}{\sigma^2}$
2. Take the derivative: $\frac{d}{dV} \ln p(D|V) = \frac{1}{\sigma^2} \sum_i \frac{1}{R_i}(I_i - \frac{V}{R_i})$
3. Find the solution by setting derivative to 0: $-\frac{1}{\sigma^2} \sum_i \frac{1}{R_i}(I_i - \frac{V}{R_i}) = 0$
$\sum_i \frac{I_i}{R_i} - V_{ML} \sum_i \frac{1}{R_i^2} = 0$
$V_{ML} = \sum \frac{I_i}{R_i} / \sum \frac{1}{R_i^2}$

2. For a given probabilistic model,
$$p(D|\theta) = b \;\; e^{-(3-\theta)^2}$$
where b is a normalising constant and a known prior of
$$p(\theta) = c \; e^{-\theta(\theta-1)}$$
where c is a normalising constant, Find the maximum a posteriori value of the model parameters θ. Clearly show the steps you followed in finding the answer.
Answer:
$$p(D|\theta)p(\theta) = be^{-(3-\theta)^2} ce^{-\theta(\theta-1)} \tag{1}$$
$$\ln p(D|\theta)p(\theta) = \ln b + \ln c - (3-\theta)^2 - \theta(\theta-1) \tag{2}$$
$$\frac{d}{d\theta} \ln p(D|\theta)p(\theta) = 2(3-\theta) - 2\theta + 1 \tag{3}$$
$$= -4\theta + 7 \tag{4}$$
$$-4\theta_{MAP} + 7 = 0 \tag{5}$$
$$\theta_{MAP} = \frac{7}{4} \tag{6}$$

3. Suppose that X is a discrete random variable with the following probability mass function, where $0 \le \theta \le 1$ is a parameter.

X	0	1	2	3
$P(X)$	$\frac{2\theta}{3}$	$\frac{\theta}{3}$	$\frac{2(1-\theta)}{3}$	$\frac{(1-\theta)}{3}$

The following 10 independent observations were taken from this distribution:

3	0	2	1	3	2	1	0	2	1

(a) What is the Maximum Likelihood estimate of θ

(b) Assume you have prior knowledge that $p(\theta) = b\,\theta\,(1-\theta)$, what would the Maximum a Posteriori (MAP) be?

Answer:

(a) Following the MLE recipe,

$$\begin{aligned} p(D|\theta) &= \prod_i P(d_i|\theta) \\ &= \left(\frac{2\theta}{3}\right)^2 \left(\frac{\theta}{3}\right)^3 \left(\frac{2(1-\theta)}{3}\right)^3 \left(\frac{1-\theta}{3}\right)^2 \\ &= c\,\theta^5\,(1-\theta)^5 (\text{where c is a constant}) \end{aligned}$$

1. Take the natural log, so

$$lnp(D|\theta) = \ln c + 5\ln\theta + 5\ln(1-\theta)$$

2. Take the derivative

$$\frac{d}{d\,\theta}\ln p(D|\theta) = \frac{5}{\theta} - \frac{5}{1-\theta}$$

3. Set the derivative to 0

$$\frac{5}{\theta_{ML}} - \frac{5}{1-\theta_{ML}} = 0$$

$$5(1-\theta_{ML}) - 5\theta_{ML} = 0$$

$$\theta_{ML} = \frac{1}{2}$$

(b) When introducing prior,

$$\begin{aligned} p(D|\theta)p(\theta) &= \left(\frac{2\theta}{3}\right)^2 \left(\frac{\theta}{3}\right)^3 \left(\frac{2(1-\theta)}{3}\right)^3 \left(\frac{1-\theta}{3}\right)^2 b\theta(1-\theta) \\ &= c\theta^6\,(1-\theta)^6 (\text{where c is a constant}) \end{aligned}$$

Following the same formula as in (a), $\theta_{MAP} = 0.5$ (same as θ_{ML} for this particular case)