Machine Learning Exercise Sheet 1

Math Refresher

Group_369

Fan Xue – fan98.xue@tum.de

Xing Zhou – xing.zhou@tum.de

Jianzhe Liu – jianzhe.liu@tum.de

November 10, 2021

Problem 8

To proof this, we need to proof a Lemma first.
Lemma:
For $a, b, c > 0$, $0 < \lambda < 1$, if $c = \lambda a + (1 - \lambda)b$, then c lies between a and b.
Proof:
Assuming that $a < b$, we have

$$c - a = (\lambda - 1)a + (1 - \lambda)b = (1 - \lambda)(b - a) > 0c - b = \lambda a - \lambda b = \lambda(a - b) < 0$$

that indicates that c lies bewteen a and b.
Now we change bach to the problem.
According to the given information, we know that the **prior distribution** $p(\theta) = Beta(a, b)$, and the **likelihood** $p(D|\theta) = Binary(m, N, \theta)$. That is

$$p(\theta) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1 - \theta)^{b-1}$$
$$p(D|\theta) = C_N^m \theta^m (1\ \theta)^{N-m}$$

We can know that the **posterior distribution**

$$p(\theta|D) \sim p(D|\theta)p(\theta) \sim \theta^{a+m-1}(1 - \theta)^{b+N-m-1}$$

Obviously the **posterior distribution** is a Beta distribution, that is

$$p(\theta|D) \sim Beta(a + m, b + N - m)$$

Let $\log p(D|\theta) = 0$, we can find the maximum likelihood estimate

$$\theta_{MLE} = \frac{m}{N}$$

From the konwledge of Beta distribution, we know the posterior mean value of θ

$$\mathbb{E}(\theta|D) = \frac{a+m}{a+b+N}$$

and the prior mean of θ

$$\mathbb{E}(\theta) = \frac{a}{a+b}$$

We can rewrite the posterior mean of θ as following

$$\mathbb{E}(\theta|D) = \frac{a+m}{a+b+N} = \frac{N}{a+b+N} \cdot \frac{m}{N} + \frac{a+b}{a+b+N} \cdot \frac{a}{a+b} = \lambda\mathbb{E}(\theta) + (1-\lambda)\theta_{MLE}$$

From the lemma we can proof that $\mathbb{E}\theta$ lies between $\mathbb{E}(\theta)$ and θ_{MLE}.