The 2nd quiz in Pattern Recognition and Machine Learning

(Answer)

3.1 Solutions

3.1.1 MLE for the Bernoulli/ binomial model

The log-likelihood is

$$\ell(\theta) = \sum_{i=1}^{N} \log \mathrm{Ber}(x_i|\theta) = \sum_{i=1}^{N} \log \left[\theta^{x_i}(1-\theta)^{1-x_i}\right] = N_1 \log\theta + N_2 \log(1-\theta) \tag{3.1}$$

where $N_1 = \sum_i x_i$ is the number of heads and $N_2 = \sum_i (1-x_i)$ is the number of tails. To find the MLE, we find the maximum of this expression as follows:

$$\frac{d\ell}{d\theta} = \frac{N_1}{\theta} - \frac{N_2}{1-\theta} = 0 \tag{3.2}$$

$$N_1 = \hat{\theta}(N_2 + N_1) \tag{3.3}$$

$$\hat{\theta} = \frac{N_1}{N_1 + N_2} \tag{3.4}$$

where $N_1 + N_2 = N$.

3.1.11 Bayesian analysis of the exponential distribution

1. The loglikelihood is

$$\ell(\theta) = N\log\theta - \theta\sum_i x_i \tag{3.37}$$

   Optimizing we get

$$\frac{d}{d\theta}\ell(\theta) = \frac{N}{\theta} - \sum_i x_i = 0 \tag{3.38}$$

$$\hat{\theta} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{3.39}$$

Wrong! The correct one is its inverse! (Mineichi)

2. $\hat{\theta}_{mle}(\mathcal{D}) = 1/\overline{x} = 1/5$.

3. $\mathbb{E}[\theta] = 1/\lambda = 1/3$ so $\hat{\lambda} = 3$

4. The prior is

$$p(\theta) = \theta^3 e^{-3\theta} \tag{3.40}$$

The likelihood is

$$p(\mathcal{D}|\theta) \propto e^{-\theta x_1} e^{-\theta x_2} e^{-\theta x_2} = e^{-15\theta} \tag{3.41}$$

So the posterior is

$$p(\theta|\mathcal{D}) \propto \theta^3 e^{-3\theta} e^{-15\theta} = \mathrm{Ga}(\theta|4, 18) \tag{3.42}$$

5. Yes, this prior is conjugate, since the exponential is a special case of the Gamma.

6. The posterior mean is

$$\frac{4}{18} = \frac{2}{9} = \frac{1}{3}\frac{3}{3+15} + \frac{1}{5}\frac{15}{3+15} \tag{3.43}$$

7. The posterior mean is a compromise between the prior mean (1/3) and the MLE (1/5). This is a more reasonable guess than the MLE since the sample size is small, so we should rely on our expert prior knowledge (although with such a simple one-parameter prior, we were not able to encode how strongly we trusted this expert).