MA2823: Foundations of Machine Learning **Homework 4**

Due 27 October, 2016.

---

1 Homework Problem

Question 1

Heights of individuals in a population have a normal distribution with unknown mean θ and standard deviation of 2. The prior distribution of θ is normal with mean $1.73m$ and standard deviation of $1m$. Twenty people are selected from the population at random and their average height is found to be $1.68m$.
(i) What is the MLE of θ and what is the prior mean?
(ii) What is the Bayes estimate of θ (using squared error loss)?
(iii) What is the limit as $n \to \infty$ and $n = 0$ of the Bayes estimator?

Solution: From the (i) The MLE of θ is the sample mean, $\hat{\theta}_{MLE} = 1.68m$, and the prior mean is $\mu = 1.73m$.
(ii) The Bayes estimate of θ is the mean of the posterior, which gives according to some precalculations ($n/\sigma_0^2 = 20/4 = 5$ and $1/\sigma^2 = 1$)

$$\begin{aligned}\theta_{Bayes} &= \frac{n/\sigma_0^2}{n/\sigma_0^2 + 1/\sigma^2}\hat{\theta}_{MLE} + \frac{1/\sigma^2}{n/\sigma_0^2 + 1/\sigma^2}\mu \\ &= \frac{5}{5+1} \cdot 1.68 + \frac{1}{5+1} \cdot 1.73 = 1.69m\end{aligned}$$

(iii) For $n \to \infty$, $\theta_{Bayes} = \hat{\theta}_{MLE} = 1.68m$. For $n = 0$, $\theta_{Bayes} = \mu = 1.73m$.

Question 2

Mark on a xy cartesian coordinate system the following five pairs
$(x, y) = \{(2, 2), (0, 0), (-2, -2), (-1, 1), (1, -1)\}$.
(i) Draw a line that you think (visually) is the best fit through those points.
(ii) Apply linear regression least-squares fit to manually compute the coefficients β of the line. What are the predicted values for y? Compare with the line you previously guessed.

Solution: (i) By visually looking at the points it is possible to think that the best fit should have slope = 1 and go through the points (-2,-2) and (2,2).
(ii) $X =$ [1 2 ; 1 0 ; 1 -2 ; 1 -1; 1 1], $y =$ [2 0 -2 1 -1]^T. Thus $(X^T X)^{-1} =$ [0.2 0; 0 0.1].

$$\hat{\beta} = (X^T X)^{-1} X^T y = [0\ 0.6]^T \tag{1}$$

Thus the predicted values for y are:

$$\hat{y} = 0.6x = [1.2\ 0\ \text{-}1.2\ \text{-}0.6\ 0.6]^T. \tag{2}$$

The visual guess is different from the calculated line. The calculated line has a slope of 0.6 and does not pass through the points (-2,-2) and (2,2).