Tutorial 13 - Preparatory Questions for the Exam

Question 1

Find the eigenvalues and eigenvectors for matrix $M = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$.

Question 2

Lucky Jim is in court accused of murder. The prosecutor in his attempt to convince the jury brings up the incredible accuracy of the DNA lab test that has been carried out. He says: "The probability of a DNA match for an innocent person is 1 in 100,000. Even more, the probability of a match for a guilty person is 1". Assume that Lucky Jim lives in a town of 105 people who could have committed the crime, and that there is a DNA match to him. The prosecutor goes on and states: "given the DNA evidence, it is very improbable that the defendant is innocent and therefore he should be found guilty". Should Lucky Jim really be declared guilty? compute the probabilities of interest and show how these change as a function of the town population (use $10^3, 10^4, ..., 10^8$).

Question 3

Consider the random variables X, Y, Z which have the following joint distribution:

$$p(X, Y, Z) = p(X)p(Y|X)p(Z|Y) \tag{1}$$

a) Show that X and Z are conditionally independent given Y.
b) if X, Y and Z are binary variables, how many parameters are needed to specify a distribution of this form?

Question 4

Given a training set consisting of four points, $([1,2],+1)$, $([3,4],+1)$, $([2,1],-1)$ and $([4,3],-1)$, find the coordinates of $\mathbf{w} = [u, v]$ and b that minimises $\|\mathbf{w}\|$ subject to

$$y_i(\mathbf{w}.\mathbf{x}_i + b) \geqslant 1, \tag{2}$$

for all $i = 1, \cdots, 4$.

Question 5

Explain why maximising the likelihood of logistic regression directly is a bad idea. What can be done to fix the problem?

Question 6

Show that in the limit $\sigma^2 \to 0$, the posterior mean for the probabilistic PCA model becomes an orthogonal projection onto the principal subspace, as in conventional PCA.

Question 7

Consider a linear model of the form

$$y(x, \mathbf{w}) = w_0 + \sum_{i=1}^{D} w_i x_i \tag{3}$$

together with a sum-of-squares error function of the form

$$E_D(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^{N} \{y(x_n, \mathbf{w}) - t_n\}^2. \tag{4}$$

Now suppose that Gaussian noise ϵ_i with zero mean and variance σ^2 is added independently to each of the input variables x_i. By making use of $E[\epsilon_i] = 0$ and $E[\epsilon_i \epsilon_j] = \delta_{ij}\sigma^2$, show that minimising E_D averaged over the noise distribution is equivalent to minimising the sum-of-squares error for noise-free input variables with the addition of a weight-decay regularisation term, in which the bias parameter w_0 is omitted from the regulariser.

学IT技匠人