VE414 Assignment 5

Sample Questions

Problem 1

Let X, y be n examples of training data and labels and let X^*, y^* be m examples of test data and labels. Let 0_n, 0_m denote zero vectors of length n, m respectively, and let k be some kernel function. Suppose that

$$p(y, y^*) = N\left(\begin{bmatrix} y \\ y^* \end{bmatrix}; \begin{bmatrix} 0_n \\ 0_m \end{bmatrix}, \begin{bmatrix} K(X,X) & K(X,X^*) \\ K(X^*,X) & K(X^*,X^*) \end{bmatrix}\right)$$

Show that the posterior distribution is

$$p(y^*|X, y, X^*) = N(y^*; \mu, \Sigma)$$

where $\mu = K(X^*, X)K(X, X)^{-1}y$ and $\Sigma = K(X^*, X^*) - K(X^*, X)K(X, X)^{-1}K(X, X^*)$.
Note: For this question, you may assume that the conditional distribution is of a Normal form, however you must derive the mean and variance.

Problem 2

Let's consider modeling f with a Gaussian process prior distribution: $p(f) = GP(f; \mu, K)$ and conditioning on the following set of data $D = (x, y)$:

$$x = [2.5, 1.5, 0.5, 0.5, 1.5, 2.5]$$

$$y = [0.0019305, 0.1054, 0.7788, 0.7788, 0.1054, 0.0019305]$$

We will fix the prior mean function μ to be identically zero; $\mu(x) = 0$.
First, let us consider about kernel selection.Consider the following two choices for the covariance function K:

$$K_1(x, x') = exp(-|x - x'|^2)$$

$$K_2(x, x') = exp(-|x - x'|)$$

Each kernel defines a Gaussian process model:

$$p(f|M_i) = GP(f; \mu, K_i)$$

We also assume a uniform prior distribution over these models:

$$P(M_i) = 1/2 \quad i = 1, 2$$

(a) Compute the log model likelihood for each model given the data D above. i.e. $\log P(D|M_i)$

(b) Compute the model posterior $P(M|D)$

(c) Please write out the predictive mean and standard deviation at $x^* = 0$ for each of the kernels, $P(y^*|x^*, D, M_i)$

Problem 3

The following table gives the number of deaths to women 80 years and older reported by day during the years 1910-1912.

Number of Deaths k	0	1	2	3	4	5	6	7	8	9
Observed Frequency n_k	162	267	271	185	111	61	27	8	3	1

(a) If the deaths are unrelated (independent) and their frequency is constant over time,a Poisson distribution $Poisson(\lambda)$ is appropriate to describe the data. What is the maximum likelihood estimation (MLE) for λ in Poisson distribution?

(b) There are likely to be different patterns of deaths in winter and summer, in which case a mixture of two Poisson distribution may provide a much better fit. Assuming a mixture of two Poisson distributions. Let p_h denote the proportion of the population from h^{th} Poisson distribution($\sum_{h=1}^{2} p_h = 1$). What is the probability of $p(y_i|\lambda, p)$?

(c) Formulate an EM algorithm for this mixture model. You need to specify the complete and the missing data and derive the E-step and the M-step.

Problem 4

Suppose we recorded number of failures Y_i and the length of operation time t_i for 20 machines: The number of failures is assumed to follow a Poisson distribution

$$Y_i \sim Poisson(\theta_i t_i), \quad i = 1, 2, \ldots 20$$

$$\theta_i \sim \Gamma(1, \beta)$$

$$\beta \sim \Gamma(1/20, 1)$$

(a) Calculate the conditional posterior distributions for θ_i and β

(b) Describe a sampler to sample the posterior distribution of θ_i and β