Department of Computer Science and Engineering
Indian Institute of Technology Madras
CS5691: Pattern Recognition and Machine Learning
Problem Sheet

Date: 18th September, 2024

1. Consider the following vector $\mathbf{x}$ and matrix $\mathbf{M}$:

   $\mathbf{x} = [1 \quad -1]^t$

   $\mathbf{M} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

   (a) Compute the value of the quadratic term of $\mathbf{x}^t\mathbf{M}\mathbf{x}$

   (b) Compute the derivative of quadratic term with $\mathbf{x}$

2. Consider the following function:

$$f(\mathbf{x}) = 2x_1^3 - 4x_2^3 + 5x_1^2x_2 - 3x_1x_2^2$$

   Give the gradient vector of $f(\mathbf{x})$ at $\mathbf{x} = [x_1 \; x_2]^t = [1 \; -1]^t$

3. A basket contains solid objects of 3 colours (Red, Green and Blue) and 2 shapes (Sphere and Cube). The number of objects of different colors and shapes are given below.

   Red Spheres: 25, Green Spheres: 15, Blue Spheres: 5

   Red Cubes: 15, Green Cubes: 10, Blue Cubes: 30

   Let the random variable X represent the color and the random variable Y represent the shape.

   (a) Give the marginal distributions $P(X)$ and $P(Y)$

   (b) Give the conditional distributions $P(X/Y = \text{Sphere})$ and $P(Y/X = \text{Green})$

   (c) Compute $P(Y = \text{Sphere}/X = \text{Green})$ using the Bayes' theorem

4. Consider the following set of 3 data points of a class:

   $\mathbf{x}_1 = [2 \quad 3]^t \qquad \mathbf{x}_2 = [1 \quad 2]^t \qquad \mathbf{x}_3 = [3 \quad 1]^t$

   Compute the covariance matrix for the data set.

5. For a bivariate Gaussian density function with $\mu = \mathbf{0}$, draw the typical level curves for each of the cases of covariance matrix given below. It may be noted that when the covariance matrix is diagonal, its eigenvectors are parallel to the coordinate axes.
   (a) Full covariance matrix with $\sigma_{12} = \sigma_{21} \neq 0$
   (b) Diagonal covariance matrix with $\sigma_1^2 \neq \sigma_2^2$
   (c) Diagonal covariance matrix with $\sigma_1^2 = \sigma_2^2 = \sigma^2$

6. For the dataset in Question 4, compute the design matrix using the polynomial basis functions of degree 2.

7. Consider the univariate data of two classes. The data of each class follows the lognormal distribution defined as follows:

$$p(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0$$

   Let the parameters of the distribution for the class i be μ_i and σ_i, with $\mu_1 \neq \mu_2$. Give the equation for the decision point of the Bayes classifier for each of the following cases:
   (a) $\sigma_1 \neq \sigma_2$ and prior probabilities of classes are not the same
   (b) $\sigma_1 = \sigma_2$ and prior probabilities of classes are the same

8. Consider a 2-class, naive-Bayes classifier for bivariate data with positive valued features x_1 and x_2 that are assumed to be statistically independent. The class-conditional density function for the feature x_j and the class i is given by an exponential density function as follows:

$$p(x_j/y_i) = \theta_{ij} e^{-\theta_{ij} x_j}$$

   Assume that the prior probabilities for the classes are the same. Show that the decision surface of the classifier is a straight line given by $w_2 * x_2 + w_1 * x_1 + w_0 = 0$. Give the expressions for w_2, w_1 and w_0.

9. For two classes, the class-conditional densities are assumed to be d-dimensional, multivariate Gaussian distributions with diagonal covariance matrices. The mean vector for the ith class is $\mu_i = [\mu_{i1}, \mu_{i2}, ..., \mu_{id}]^t$. The covariance matrix for the ith class is $C_i = \text{diag}(\sigma_{i1}^2, \sigma_{i2}^2, ..., \sigma_{id}^2)$. Then the squared Mahalanobis distance

$$(\mathbf{x} - \mu_i)^t C_i^{-1} (\mathbf{x} - \mu_i) = \sum_{j=1}^{d} \frac{(x_j - \mu_{ij})^2}{\sigma_{ij}^2}$$

Assume that the prior probabilities for the classes are equal. Show that the equation for the decision surface of the Bayes classifier for the two classes is given by a quadratic equation

$$g_{12}(\mathbf{x}) = \sum_{j=1}^{d} a_2 * x_j^2 + a_1 * x_j + a_0 = 0$$

Give the expressions for coefficients a_2, a_1 and a_0.

10. The class-conditional density for a class follows the Maxwell distribution defined as follows:

$$p(x/\theta) = \frac{4}{\sqrt{\pi}} \theta^{\frac{3}{2}} x^2 e^{-\theta x^2}, \text{for} \quad x \geq 0$$

Give the maximum likelihood estimate of the parameter θ using the training data set $D = \{x_1, x_2, ..., x_N\}$.