Subject Name: Statistical Foundation of Machine Learning

Mid-sem Exam, 30 Points, Time : 2 hrs exam. (All questions are compulsory)

Q1 : (10 points) :

In the below joint probability density, the three data points are observed. Kindly find out the values of μ and σ that results in giving the maximum value of the below expression.

$$P(9, 9.5, 11; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(9-\mu)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(9.5-\mu)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(11-\mu)^2}{2\sigma^2}\right)$$

Q2 : (5 points)

$X = (X_1, X_2)$ is drawn from a two dimensional Gaussian distribution with a

diagonal covariance matrix.

$$X = (X_1, X_2) \sim \mathcal{N}(\mu, \Sigma)$$
$$\Sigma = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

where a and b are some real numbers.

Are X_1 and X_2 independent? Explain as succinctly as possible.

Q3 : (5 points) :

Prove that the eigenvalues in eigenvalue decomposition are always real. Further when they are distinct, the eigenvectors should be orthogonal.

Q4 : (6 points)

(6 points) We are given a set of two dimensional inputs and their corresponding output pair: $\{x_{i,1}, x_{i,2}, y_i\}$. We would like to use the following regression model to predict y:

$$y_i = w_1^2 x_{i,1} + w_2^2 x_{i,2}$$

Derive the optimal value for w_1 when using least squares as the target minimization function (w_2 may appear in your resulting equation). Note that there may be more than one possible value for w_1.

Q5 : (4 points) : Prove two similar matrices A and B have the same characteristic polynomial hence same eigenvalues.