March 7, 2022

SML - Mid Sem Winter 2022

Note:

- Symbols have their usual meanings. Small letter in bold font indicates vector. Capital letter in bold font indicates matrix. Otherwise, the dimensions should be clear from the context.
- Duration 2 hours.
- Number in [.] indicate marks
- If you have any extra variable/constant in the expression, do clearly define it.

Q1. Let $\mathbf{x}$ be a d-dimensional random vector and distributed according to some arbitrary distribution with mean $\boldsymbol{\mu}$ and covariance $\mathbf{S}$. Let $y = \mathbf{w}^\top \mathbf{x} + b$. Suppose you have N iid samples from the distribution $p(y) = Ce^{-(y-\mu_y)^2}$, where C is a scalar constant. Find MLE estimate of $\mathbf{w}$, such that the constraint $\mu_y^2 = \sigma_y^2$ is satisfied. $\mathbf{w}$ should be expressed in terms of N, $\boldsymbol{\mu}$, $\mathbf{S}$, b, y. [4]

Q2. Suppose there is a system that uses FDA before binary classification. Assume that you have complete access to the system but your role is that of an attacker. As an attacker, you would want that the classification performance is as poor as possible. Since you know that the system applies FDA, you decide to find a vector $\mathbf{w}$ which is obtained in the following way. Given the data matrices for two classes $\mathbf{X}_1$ and $\mathbf{X}_2$, first project them using $\mathbf{w}$. Assume the number of samples in each class to be $N/2$. Then, assume that the projected samples (including both the classes that is all N samples) follow a Bernoulli distribution with known parameter θ as well as these samples are iids. Further, let' say that $\mathbf{w}$ is drawn from a multivariate Gaussian with zero mean and identity covariance with the elements of $\mathbf{w}$ being statistically independent. In order to attack the system, a constraint is also introduced which is $\mathbf{w}^\top \boldsymbol{\mu}_1 = \mathbf{w}^\top \boldsymbol{\mu}_2$, where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are the respective means of the two classes. Find an expression for $\mathbf{w}$. Express $\mathbf{w}$ in terms of $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, θ, N. [4]

Q3. For a regression setting, let the data matrix be $\mathbf{X}$ and the labels are $\mathbf{Y}$. Suppose we want the estimate of mean of the error values $\mathbf{e} = \mathbf{Y} - \mathbf{W}^\top \mathbf{X}$. Let all the dimensions of $\mathbf{e}$ be statistically independent and distributed according to multivariate Gaussian with unknown mean $\boldsymbol{\mu}$ and a known covariance $\mathbf{S}$. Using a prior $\lambda e^{-\|\boldsymbol{\mu}\|_2^2}$, find MAP estimate of $\boldsymbol{\mu}$ in terms of $\mathbf{S}$, $\mathbf{w}$, N (number of iid samples), $\boldsymbol{\mu}_x$ (mean of $\mathbf{X}$), $\boldsymbol{\mu}_y$ (mean of $\mathbf{Y}$). [3]

Q4. Let

$$\begin{aligned} p(\mathbf{x}|\omega_1) &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\ p(\mathbf{x}|\omega_2) &\sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{I}) \\ p(\mathbf{x}|\omega_3) &\sim \mathcal{N}(-\boldsymbol{\mu}, \mathbf{I}) \end{aligned}$$

where $\mathbf{I}$ is 2×2 identity matrix and $\mathbf{x}$ is a 2-dimensional vector. Let us project $\mathbf{x}$ using $\mathbf{u}$. Let the projections be called y. We want that the probability of projections belonging to respective classes be maximum. This mean we have to perform

$$\max_{\mathbf{u}} \; p(y|\omega_1; \mathbf{x}) p(y|\omega_2; \mathbf{x}) p(y|\omega_3; \mathbf{x}) \tag{1}$$

$$\begin{aligned} p(y|\omega_1; x) &= \frac{1}{Z} exp\{-.5y^\top y\} \\ p(y|\omega_2; x) &= \frac{1}{Z} exp\{-.5(y - \mathbf{u}^\top\boldsymbol{\mu})^\top (y - \mathbf{u}^\top\boldsymbol{\mu})\} \\ p(y|\omega_3; x) &= \frac{1}{Z} exp\{-.5(y + \mathbf{u}^\top\boldsymbol{\mu})^\top (y + \mathbf{u}^\top\boldsymbol{\mu})\} \end{aligned}$$

Z is a normalization constant. $p(y|\omega_i) \quad i = 1, 2, 3$, denotes the probability of projected encodings of $\mathbf{x}$ conditioned on respective classes. However, this does not capture FDA assumptions which can help in better discrimination. Using Equation 1 and knowledge of FDA, find an expression for $\mathbf{u}$. Note you need to create a function of $\mathbf{u}$ which has both Equation 1 and objective of FDA. Assume the total scatter matrix is known to be $\mathbf{S}$. $\mathbf{u}$ should be expressed in terms of $\mathbf{S}$, $\boldsymbol{\mu}$, number of samples in each class N. [4]