Indian Institute of Technology Jodhpur

Machine Learning I: Minor-2 Exam

Date: May 1, 2021, Max Marks: 20

Max Time: 60 min (extra 10 min for submission)

`Attempt all questions. Best of luck` ☺

1. (a) Consider two unit norm vectors $\mathbf{x}$ and $\mathbf{y}$. Find the distance between $\mathbf{x}$ and $\mathbf{y}$ if the angle between these two vectors is $45°$. **[3 Marks]**

   (b) Let $\mathbf{A} \in \mathbb{R}^{3\times 3}$ be a matrix such that $\mathbf{A}(i,j) = ij$ for all $i, j = 1, 2, 3$. Let $\mathbf{v}_1$ be the eigenvector of $\mathbf{A}$ corresponding to the largest eigenvalue and let $\mathbf{v}_3$ be the eigenvector of $\mathbf{A}$ corresponding to the smallest eigenvalue. Find the value of the quantity $16\mathbf{v}_1^\top \mathbf{v}_3$. **[2 Marks]**

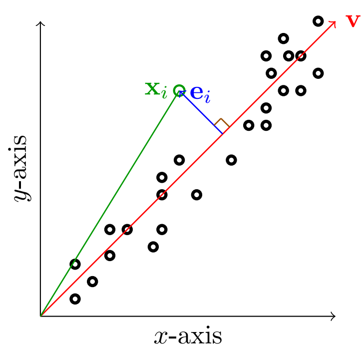
2. Consider a coin with unknown probability of head. Let X be a random variable taking value 1 if head appears and 0 if tail appears. Let us assume that we are given a set of 10 independent samples $\{1, 1, 0, 0, 0, 1, 1, 1, 0, 1\}$. Assume that the probability mass function of X is parameterized by θ and is defined as $p_\theta(x) = \theta^x(1-\theta)^{1-x}$. Find the maximum likelihood estimator value of θ. **[5 Marks]**

3. The principal component analysis algorithm can be used to find the direction of maximum variance for a given set of points. For example, consider the data points given in below figure. We can use the PCA algorithm to find the direction (shown as a red colored vector) of maximum variance $\mathbf{v}$. Now, let us formally consider a set $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ of n mean centered points in $\mathbb{R}^2$. Let $\mathbf{X} = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n\end{bmatrix} \in \mathbb{R}^{2\times n}$ be a matrix containing these points. Assume that $\mathbf{v}$ is a unit norm vector. Complete the following steps.

   (a) Show that $\mathbf{e}_i = (\mathbf{I} - \mathbf{v}\mathbf{v}^\top)\mathbf{x}_i$. **[1.5 Marks]**

   (b) Show that $\underset{\mathbf{v}\in\mathbb{R}^2, \mathbf{v}^\top\mathbf{v}=1}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n} \|\mathbf{e}_i\|_2^2 = \underset{\mathbf{v}\in\mathbb{R}^2, \mathbf{v}^\top\mathbf{v}=1}{\operatorname{argmax}} \frac{1}{n}\mathbf{v}^\top\mathbf{X}\mathbf{X}^\top\mathbf{v}$. **[2.5 Marks]**

   (c) Show that best $\mathbf{v}$ is the eigenvector corresponding to the largest eigenvalue of $\frac{1}{n}\mathbf{X}\mathbf{X}^\top$. **[1 Mark]**



4. Consider a set $\mathcal{S} = \{1, 2, 9, 10\}$ of four independent samples drawn from a mixture of two Gaussian distributions $\mathcal{N}(\mu_1, \sigma_1^2)$ and $\mathcal{N}(\mu_2, \sigma_2^2)$ with mixture weights π_1 and π_2. Consider the initial values of these parameters as $\mu_1 = 0$, $\mu_2 = 7$, $\sigma_1^2 = \sigma_2^2 = 1$ and $\pi_1 = \pi_2 = \frac{1}{2}$. Use the expectation maximization algorithm to estimate the parameters of the first Gaussian distributions. First find the responsibilities $\gamma(jt)$ given these initial parameters, here, $j = 1, 2, 3, 4$ and $t = 1$. Then, update the parameters μ_1, σ_1^2, and π_1 using the calculated responsibilities. **[5 Marks]**