CSE 326/426 (Fall 2022) Homework 7

Due on 11:55pm, Fri, Oct 28, 2022

Grading: All questions have the same points (10 each). Only questions with $*$ will be graded (undergraduates: do not need to answer the graduate-only question).

Submitting: Only electronic submissions in PDF format on Coursesite are accepted. Name your file as

<Your LIN>_HW7.pdf

A few ways to create a single PDF file. i) Use Microsoft Word and insert your writing or any figure taken, scanned, or plotted. Then choose "File→Print" in the main menu and you will find an option for outputting the file to a PDF file. ii) Use Latex to write your solution and include any figures. iii) Use the online Google Doc as an alternative of Word. It has sufficient features to combine multiple images and texts. Exporting to PDF files is similar as with Word.

Please DO NOT compress the PDF file as that will slow down the processing of your submission and the grading.

Questions:

1. $*$ Prove that if $p(x)/q(x) = c$ for two probability density functions p and q over the same domain for random variable x, then $c = 1$.

2. Create two Bernoulli distributions $p(X)$ and $q(X)$ of a random variable $X \in \{0, 1\}$ so that $p \neq q$ but

$$\mathrm{KL}(p \parallel q) = \mathrm{KL}(q \parallel p). \tag{1}$$

   The KL-divergence between two distributions of a discrete random varible $X \in \{x_1, \dots, x_K\}$ is defined as

$$\mathrm{KL}(p \parallel q) = \sum_{k=1}^{K} p(X = x_k) \log \frac{p(X = x_k)}{q(X = x_k)}. \tag{2}$$

3. $*$ Prove the following equality about the squared error of approximating the vector $\mathbf{x} \in \mathbb{R}^n$ using $z_1\mathbf{u}_1 + z_2\mathbf{u}_2$ where $\mathbf{u}_1$ and $\mathbf{u}_2$ are two eigenvectors in PCA and z_1 and z_2 are the two corresponding coefficients so that $z_j = \mathbf{x}^\top \mathbf{u}_j$ for $j = 1, 2$.

$$\|\mathbf{x} - z_1\mathbf{u}_1 - z_2\mathbf{u}_2\|^2 = \|\mathbf{x}\|^2 - \sum_{j=1}^{2} z_j^2. \tag{3}$$

   ($\|\mathbf{x}\|^2 = \mathbf{x}^\top\mathbf{x}$ is the square of the Euclidean norm of the vector $\mathbf{x}$.)

4. Prove that the projection $B(B^\top B)^{-1}B^\top\mathbf{x}$ is a project of $\mathbf{x}$. (*Hints: try to apply the same projection to the vector* $B(B^\top B)^{-1}B^\top\mathbf{x}$ *and show that the result remains* $B(B^\top B)^{-1}B^\top\mathbf{x}$.)

5. $*$ (Graduate only) List one commonality and one difference between LDA (linear discriminant analysis) and SVM.