ELE 535 - Machine Learning and Pattern Recognition

Fall 2018

HOMEWORK 1: Theory

Q1 Consider the training data $\{(x_j, y_j)\}_{j=1}^m$, with $x_j \in \mathbb{R}^n$ and $y_j \in \{0, 1\}$. Assume the training data has an equal number of examples from each class. Hence the estimated prior probabilities of each class are equal. The nearest centroid classifier has

$$\hat{y}(x) = \begin{cases} 1, & \text{if } \|x - \hat{\mu}_1\|_2 < \|x - \hat{\mu}_0\|_2; \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that the nearest centroid classifier is a linear classifier with

$$\begin{aligned} w &= (\hat{\mu}_0 - \hat{\mu}_1) \\ b &= (\hat{\mu}_1 - \hat{\mu}_0)^T \frac{\hat{\mu}_1 + \hat{\mu}_0}{2} \end{aligned}$$

(b) Show that the classifier can also be written as

$$\hat{y}(x) = \begin{cases} 1, & \text{if } <\hat{\mu}_0 - \hat{\mu}_1, x - \hat{\mu}> < 0; \\ 0, & \text{otherwise.} \end{cases}$$

Here $\hat{\mu} \triangleq \frac{1}{m} \sum_{i=1}^m x_i$ denotes the mean of the training examples. So the result of classification depends solely on the sign of a inner product.

(c) By neatly sketching the vectors $x - \hat{\mu}$ and $\hat{\mu}_0 - \hat{\mu}_1$, give a geometric interpretation of this classifier.

(d) Suppose we first center the training data by subtracting $\hat{\mu}$ from each training example. Determine the form of the nearest centroid classifier for the centered training data.

Q2 For a $n \times m$ real matrix X show that:

(a) $\mathcal{R}(X) \triangleq \{z \in \mathbb{R}^n : z = Xw, \text{for } w \in \mathbb{R}^m\}$ is a subspace of $\mathbb{R}^n$.

(b) $\mathcal{N}(X) \triangleq \{a \in \mathbb{R}^m : Xa = \mathbf{0}\}$ is a subspace of $\mathbb{R}^m$.

Q3 (a) Let $A_j \in \mathbb{R}^{n_j \times m}$ and $\mathcal{N}_j = \{x \in \mathbb{R}^m : A_j x = \mathbf{0}\}$, $j = 1, 2$. Show that $\mathcal{N}_1 \cap \mathcal{N}_2$ is a subspace of $\mathbb{R}^m$. Give a similar matrix equation for this subspace.

(b) Let $A_j \in \mathbb{R}^{n \times m_j}$ and $\mathcal{R}_j = \{y \in \mathbb{R}^n : y = A_j x \text{ with } x \in \mathbb{R}^{m_j}\}$, $j = 1, 2$. Show that $\mathcal{R}_1 + \mathcal{R}_2 = \{y_1 + y_2 : y_1 \in \mathcal{R}_1, y_2 \in \mathcal{R}_2\}$ is a subspace of $\mathbb{R}^n$. Give a similar matrix equation for the subspace.

Q4 For $A, B \in \mathbb{R}^{m \times n}$, the inner product of A and B is $<A, B> \triangleq \sum_i \sum_j A_{i,j} B_{i,j}$. Show that $<A, B> = \operatorname{trace}(A^T B)$.

Q5 Consider the vector space of real $n \times n$ matrices. Let $\mathcal{S}$ and $\mathcal{A}$ denote the subsets of symmetric ($P^T = P$) and antisymmetric ($A^T = -A$) matrices, respectively. Show that $\mathcal{S}$ and $\mathcal{A}$ are subspaces of $\mathbb{R}^{n \times n}$ and that $\mathcal{S}^\perp = \mathcal{A}$. Hence $\mathbb{R}^{n \times n} = \mathcal{S} \oplus \mathcal{A}$.