ELE 535: Machine Learning and Pattern Recognition
Homework 3

Zachary Hervieux-Moore

Monday 8th October, 2018

Exercise 1: In Chapter 1 we developed the binary MAP classifier

$$f(x) = \begin{cases} 1, & \text{if } \ln\left(\frac{p_1(x)}{p_0(x)}\right) > \ln\left(\frac{p(0)}{p(1)}\right) \\ 0, & \text{otherwise} \end{cases}$$

This is based on an underlying generative model in which $p(k)$ is the prior probability of class k, and $p_k(x) = p(x|k)$ the conditional density of the datum x given that the class k, $k \in \{0, 1\}$. Now assume that $x \in \mathbb{R}^n$, and that $p_k(x)$ is a multivariate Gaussian density

$$p_k(x) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma_k|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)}$$

a) Determine the resulting MAP classifier in its simplest form.

b) Determine the form of the decision boundary for this classifier.

c) An empirical version of the MAP classifier is obtained by using training data to estimate any unknown parameters. Then using these estimates in the MAP classifier. Compare this empirical Bayes classifier to the nearest centroid classifier.

Answer:

Exercise 2: (Naive Bayes classifier) Derive the MAP classifier when the conditional probability density $p_k(x)$ is the multivariate Gaussian density as in 1) with Σ_k a diagonal matrix, $k = 0, 1$. As before, the prior probability $p(k)$ of class $k \in \{0, 1\}$ is given

a) Determine the resulting MAP classifier in its simplest form.

b) Determine the form of the decision boundary for this classifier.

c) An empirical version of the MAP classifier is obtained by using training data to estimate any unknown parameters. Then using these estimates in the MAP classifier. This yields the Naive Bayes classifier.

Answer:

Exercise 3: Let $X \in \mathbb{R}^{n\times m}$ be a data matrix with data items stored in the columns of X. Show that the set of nonzero eigenvalues of XX^T is the same as the set of nonzero eigenvalues of X^TX.

Answer:

Exercise 4: Some additional properties of singular values. Let $\sigma_i(A)$ denote the i^{th} singular value of the matrix A, $i \in [1 : r]$ with $r = \text{rank}(A)$. Prove the following.

a) If λ is an eigenvalue of $A \in \mathbb{R}^{n \times n}$, then $|\lambda| \le \sigma_1(A)$.

b) For $A \in \mathbb{R}^{m \times n}$, $|A_{i,j}| \le \sigma_1(A)$, $i \in [1 : m]$, $j \in [1 : n]$.

c) For $\alpha \in \mathbb{R}$ and $A \in \mathbb{R}^{m \times n}$, $\sigma_i(\alpha A) = |\alpha| \sigma_i(A)$, $i \in [1 : r]$.

Answer:

Exercise 5: Nuclear Norm. Let $A \in \mathbb{R}^{m\times n}$ have rank r with $r \leq q = \min(m,n)$. Define the nuclear norm of A by $\|A\|_* = \sum_{i=1}^{r} \sigma_i(A)$.

a) Find $B \in \mathbb{R}^{m\times n}$ that maximizes $\langle A, B\rangle$ subject to $\sigma_1(B) \leq 1$.

b) Show that $\|A\|_* = \max_{\|C\|_2 \leq 1} \langle A, C\rangle$.

c) Show that $\|\cdot\|_*$ is a norm on $\mathbb{R}^{m\times n}$.

Answer: