ECE 175A - Homework Set Three

Due Thursday, January 30, 2020 (lecture 8)

1. Suppose that we have a classification problem with two classes of equal probability, i.e. $P_Y(0) = P_Y(1) = 1/2$, and class-conditional densities of the form

$$p_{X|Y}(x|j) = K_j e^{-\frac{|x-a_j|}{b_j}}, \ j = 0, 1, \ b_j > 0.$$

a) Determine the value of the constants K_j.

b) Calculate the Bayes decision rule for this classifier, as a function of the parameters a_j and b_j.

c) For the case where $a_0 = 0$, $b_0 = 1$, $a_1 = 1$, and $b_1 = 2$, what is the range of x that are classified with the label 0?

d) Repeat **c)** for the case in which $P_Y(0) = 0.75$.

2. Consider a classification problem with two Gaussian classes of identical covariance $\mathbf{\Sigma} = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, i.e.

$$p_{\mathbf{X}|Y}(\mathbf{x}|j) = \frac{1}{\sqrt{2\pi|\sigma^2\mathbf{I}|}} \exp\{-\frac{1}{2\sigma^2}(\mathbf{x}-\mu_j)^T(\mathbf{x}-\mu_j)\}$$

a) Show that, in this case, the optimal decision boundary is an hyper-plane by showing that the set of points in the boundary, i.e. the points $\mathbf{x}$ such that

$$P_{\mathbf{X}|Y}(\mathbf{x}|0)P_Y(0) = P_{\mathbf{X}|Y}(\mathbf{x}|1)P_Y(1),$$

satisfy the hyper-plane equation

$$\mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) = 0.$$

Determine the values of $\mathbf{w}$ and $\mathbf{x}_0$ as a function of the priors $P_Y(j)$, the Gaussian means μ_j and the variance σ^2.

b) Consider the 2D case case in which $\mu_0 = (-1, -1)$, $\mu_1 = (1, 1)$, $\sigma^2 = 1/4$, and $P_Y(1) = 0.5$. Using MATLAB, make a plot showing a few contours of each Gaussian, and the line corresponding to the optimal decision boundary (note: to draw the Gaussian contours, create an x, y grid using `meshgrid()`, evaluate the values of Gaussian associated with each class at each point of this grid, multiply by the priors, add the two scaled Gaussians, and make a contour plot using `contour()`). Repeat for $P_Y(1) = 0.25$ and $P_Y(1) = 0.75$.

c) Can you give a geometric interpretation to the vector $\mathbf{w}$ and the point $\mathbf{x}_0$?

3.(Computer) This week we will make a move towards learning a distribution for our classes. We will assume the class-conditional densities have a multivariate Gaussian distribution of $28 \times 28 = 784$ dimensions

$$P_{X|Y}(x|i) = \frac{1}{\sqrt{(2\pi)^n|\Sigma_i|}} \exp\left\{-\frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)\right\}$$

Remember from our previous experiment that the feature space was $\mathbb{R}^{28\times 28}$.

Moreover, we shall assume that the covariance matrix for every class is the same and equal to identity covariance, i.e. $\Sigma_i = \Sigma = \mathbf{I}$, $\forall\ i$ and also that the prior for the classes are distributed uniformly, i.e $P_Y(i) = 1/N =$ constant where N is the number of classes. We shall use the same database as used in previous assignment.

1. For each digit, calculate and display the sample mean (See Notes)
2. Using sample mean as an estimate of the class mean (μ_i), perform the task of classification by Bayes Decision Rule.

$$i^*(x) = argmax_i \left\{ -\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i) - \frac{1}{2}\log(2\pi)^d|\Sigma_i| + \log P_Y(i) \right\}$$

   (a) Again for each class calculate and plot the error rates, P(Error|Y = i) for all $i = 0 \ldots 9$.
   (b) Calculate the total error rate, P(Error).
3. **(Optional, extra credit)**. Now for different digits, calculate and display the covariance matrix, you can use `cov` function to calculate the 784×784 dimensional covariance matrix and `imshow(imatrix,[]);` to display. Why were we not able to use this matrix in our distance computations?

- You should use subplot to display multiple figures in the same window
- Once you calculate and display the mean, its a good idea to use the 'stacked' version for mean and test images. By stacked, we mean to convert a $n \times n$ matrix to a $n^2 \times 1$ vector. In matlab you can use the `reshape` command or $M_{stacked} = M(:)$