Machine Learning & Machine Learning (extended)

Practice Exercise Sheet – Bayesian and Covariance

Question 1: What is the effect of covariance between attributes and the resultant shape modelled for the point cloud (i.e. the training samples within a class)? You can study the effect under various scenarios: (i) increase in covariance, (ii) decrease in covariance, (iii) positive covariance, (iv) negative covariance, (v) zero covariance, (vi) identical covariance for all classes, (vii) arbitrary (i.e. non-identical) covariance for different classes, etc. You may find the following MATLAB's functions useful to determine the covariance between attributes and studying the effect of covariance on the shape of point cloud: cov, mvnrnd, mvnpdf.

Question 2: Consider a 2 class Bayesian classifier which is trained from the past examples to predict the target labels by computing the posterior estimates. However, this classifier can be considered as a model that draws a decision boundary (linear or non-linear) between the classes such that this boundary separates the training samples. Answer the following questions about the Gaussian classifier:

(i) Is it possible for a Gaussian classifier to implement a non-linear decision boundary? If so, draw an example and suggest the shape of this non-linear decision boundary. If not, explain why not.

(ii) How about a Gaussian Naive Bayes classifier? Justify your answer.

Hints:

With multivariate Gaussian, the following cases can be considered:

a) statistically independent attributes, identically distributed Gaussian for each class (i.e. same variance for each class, and 0 entries at the non-diagonal location in the covariance matrix)

b) identical covariance for each class (i.e. $\Sigma = \Sigma_1 = \Sigma_2$)

c) arbitrary (non-identical) covariance for each class (i.e. $\Sigma_1 \neq \Sigma_2$)

There are two ways to attempt this problem.

In an informal way, you can attempt this problem by considering the shape of each class (for example by drawing the density contours around training samples) under the influence of covariance matrix.

In a more formal (and mathematical way), you can consider the maximum a posteriori estimate and determining the shape of the decision boundary in each of the above cases.

Question 3: Let's consider the mean and covariance of two class dataset below.

$\boldsymbol{\mu}_1 = \begin{bmatrix}2\\2\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}, \boldsymbol{\mu}_2 = \begin{bmatrix}8\\8\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}$

Can you make a rough drawing of how the point cloud of each class would be modelled with multivariate Gaussian pdf? What will be the shape of the decision boundary with a Gaussian classifier, in the cases of with and without naïve assumption?

Question 4: Repeat the above with the following mean and covariance of two class dataset.

$\boldsymbol{\mu}_1 = \begin{bmatrix}2\\2\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}, \boldsymbol{\mu}_2 = \begin{bmatrix}8\\8\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & 0\\0 & 3\end{bmatrix}$

Question 5: Repeat the above with the following mean and covariance of two class dataset.

$\boldsymbol{\mu}_1 = \begin{bmatrix}2\\2\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & 2\\2 & 3\end{bmatrix}, \boldsymbol{\mu}_2 = \begin{bmatrix}2\\2\end{bmatrix}, \boldsymbol{\Sigma}_1 = \begin{bmatrix}2 & -2\\-2 & 4\end{bmatrix}$

Question 6: Consider a 2-class problem. Let $p(\boldsymbol{x}|c)$, where $c \in (1,2)$, be a multivariate Gaussian pdf, with mean vector $\boldsymbol{\mu}_c$, and covariance matrix $\boldsymbol{\Sigma}_c$. Let us assume that the covariance matrix is same for both classes with same variance for all attributes and the non-diagonal entries are zero. For this case, show with a graphical illustration the modelled shape of point cloud of each class and the shape of the decision boundary of the Gaussian classifier.

Question 7: Is it possible for a Gaussian classifier to implement a non-linear decision boundary? If so, draw an example. If not, explain why not.

Question 8: Is it possible for a Gaussian Naïve Bayes classifier to implement a non-linear decision boundary? If so, draw an example. If not, explain why not.

(Bonus) Question: [Optional] Repeat the above to mathematically show whether it has the form of a linear boundary $\boldsymbol{w}^T\boldsymbol{x} + b$ with some weight vector $\boldsymbol{w}$ and some scalar b.