ECE M146 Discussion 6

Introduction to Machine Learning

Instructor: Lara Dolecek

TA: Zehui (Alex) Chen, Ruiyi (John) Wu

1. More about Discriminative v.s. Generative

   Let $p(x|C_1) \sim 0.4\mathbb{N}(0.2, 0.1) + 0.6\mathbb{N}(0.5, 0.1)$. Let $p(x|C_2) \sim \mathbb{N}(0.7, 0.1)$. In MATLAB, plot the two class conditional distribution and find the decision boundary. Let $P(C_1) = P(C_2) = 0.5$, what is the equation to find the posterior distribution for C_1 and C_2. Find and plot the posterior distribution for C_1 and C_2.

   Find the maximum likelihood decision boundary using both the class conditional distribution and the posterior distribution. Comment on your observation.

2. We are given a training set $\{(x^{(i)}, y^{(i)}); i = \{1, \cdots, m\}\}$, where $x^{(i)} \in R^n$ and $y^{(i)} \in \{0, 1\}$. We consider the Gaussian Discriminant Analysis (GDA) model, which models $P(x|y)$ using multivariate Gaussian. Writing out the model, we have:

$$P(y = 1) = \phi = 1 - P(y = 0)$$

$$P(x|y = 0) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_0)^T \Sigma^{-1} (x - \mu_0)\right)$$

$$P(x|y = 1) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu_1)^T \Sigma^{-1} (x - \mu_1)\right)$$

The log-likelihood of the data is given by:

$$L(\phi, \mu_0, \mu_1, \Sigma) = \ln P(x^{(i)}, \cdots, x^{(m)}, y^{(i)}, \cdots, y^{(m)}) = \ln \prod_{i=1}^{m} P(x^{(i)}|y^{(i)}) P(y^{(i)}).$$

In this exercise, suppose we already find μ_0 and μ_1, we want to maximize $L(\phi, \mu_0, \mu_1, \Sigma)$ with respect to Σ.

(a) Write down the explicit expression for $P(x^{(i)}, \cdots, x^{(m)}, y^{(i)}, \cdots, y^{(m)})$ and $L(\phi, \mu_0, \mu_1, \Sigma)$.

(b) Differentiate $L(\phi, \mu_0, \mu_1, \Sigma)$ with respect to Σ and set it to 0. Show that the maximum likelihood result for Σ is:

$$\Sigma = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)} - \mu_{y^{(i)}})(x^{(i)} - \mu_{y^{(i)}})^T.$$

Hints: You may use the following properties without proof: $a = Tr(a)$ for scalar a; $Tr(A) + Tr(B) = Tr(A + B)$; $\frac{\partial \ln |A|}{\partial A} = A^{-T}$; $\frac{\partial Tr(A^{-1}B)}{\partial A} = -(A^{-1}BA^{-1})^T$.