Assignment Two

Matthew C. Scicluna
Département d'Informatique et de Recherche Opérationnelle
Université de Montréal
Montréal, QC H3T 1J4
matthew.scicluna@umontreal.ca

October 5, 2017

1 Fisher LDA

Given the class variable, the data are assumed to be Gaussians with different means for different classes but with the same covariance matrix.

1.1 Derive the form of the maximum likelihood estimator for this model

Given $Y \sim Bernoulli(\pi), X \mid Y = j \sim \mathcal{N}(\mu_j, \Sigma)$. We first obtain the log likelihood $l(\theta \mid D)$, where D are the data points $\{x^{(i)}, y^{(i)}\}_{i=1}^N$ and $\theta = (\pi, \mu_0, \mu_1, \Sigma)$, $\pi \in [0, 1]$, $\mu_0, \mu_1 \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d \times d}$.

$$
\begin{aligned}
l(\theta \mid D) &= \ln P(D \mid \theta) \\
&= \sum_{i=1}^{N} \ln P(x^{(i)}, y^{(i)} \mid \theta) \\
&= \sum_{i=1}^{N} \ln P(x^{(i)} \mid y^{(i)} \mid \theta) + \ln P(y^{(i)} \mid \theta) \\
&\propto \sum_{i=1}^{N} -\frac{1}{2} \ln |\Sigma| - \frac{1}{2}(x^{(i)} - \mu_{y^{(i)}})^T \Sigma^{-1} (x^{(i)} - \mu_{y^{(i)}}) + y^{(i)} \ln \pi + (1 - y^{(i)}) \ln(1 - \pi)
\end{aligned}
$$