# CSC411 Fall 2017
# Assignment 2 Report

Tianbao Li

2017/11/04

## 1 Q1: Class-conditional gaussians

Given:

$$p(y = k) = a_k \tag{1}$$

$$p(\mathbf{x}|y = k, \mu, \sigma) = (\prod_{i=1}^{d} 2\pi\sigma_i^2)^{-1/2} \exp\{-\sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i - \mu_{ki})^2\} \tag{2}$$

### 1.1 Bayes' rule derivation

$$\begin{aligned}
p(y = k|\mathbf{x}, \mu, \sigma) &= \frac{p(\mathbf{x}|y = k, \mu, \sigma)p(y = k)}{p(\mathbf{x}|\mu, \sigma)} \\
&= \frac{p(\mathbf{x}|y = k, \mu, \sigma)p(y = k)}{\sum_{j=1}^{K} p(\mathbf{x}|y = j, \mu, \sigma)} \\
&= \frac{(\prod_{i=1}^{d} 2\pi\sigma_i^2)^{-1/2} \exp\{-\sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i - \mu_{ki})^2\}a_k}{\sum_{j=1}^{K}(\prod_{i=1}^{d} 2\pi\sigma_i^2)^{-1/2} \exp\{-\sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i - \mu_{ji})^2\}}
\end{aligned}$$

### 1.2 Negative likelihood function (NLL)

$$\begin{aligned}
\ell(\theta; D) &= -\log p(y^{(1)}, \mathbf{x}^{(1)}, y^{(2)}, \mathbf{x}^{(2)}, \dots, y^{(N)}, \mathbf{x}^{(N)}|\theta) \\
&= -\log \prod_{n=1}^{N} p(y^{(n)}, x^{(n)}|\theta) \\
&= -\sum_{n=1}^{N} (\log p(x^{(n)}|y^{(n)}, \theta) + \log p(y^{(n)}|\theta)) \\
&= -\sum_{n=1}^{N} (\log((\prod_{i=1}^{d} 2\pi\sigma_i^2)^{-1/2} \exp\{-\sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i^{(n)} - \mu_{ki})^2\}) + \log\alpha_k) \\
&= -\sum_{n=1}^{N} (-\frac{1}{2}\sum_{i=1}^{d} 2\pi\sigma_i^2 - \sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i^{(n)} - \mu_{ki})^2 + \log\alpha_k) \\
&= \sum_{n=1}^{N} (\sum_{i=1}^{d} \pi\sigma_i^2 + \sum_{i=1}^{d} \frac{1}{2\sigma_i^2}(x_i^{(n)} - \mu_{ki})^2 - \log\alpha_k)
\end{aligned}$$

### 1.3 Partial derivatives

$$
\begin{aligned}
\frac{\partial \ell}{\partial \mu_{ki}} &= \frac{\partial(\sum_{n=1}^{N}\sum_{i=1}^{d}\frac{1}{2\sigma_i^2}(x_i^{(n)}-\mu_{ki})^2)}{\partial \mu_{ki}} \\
&= -\sum_{n=1}^{N}\sum_{i=1}^{d}\frac{1}{\sigma_i^2}\mathbb{1}(y^{(n)}=k)(x_i^{(n)}-\mu_{ki}) \\
\frac{\partial \ell}{\partial \sigma_i^2} &= \frac{\partial(\sum_{n=1}^{N}(\sum_{i=1}^{d}\pi\sigma_i^2+\sum_{i=1}^{d}\frac{1}{2\sigma_i^2}(x_i^{(n)}-\mu_{ki})^2))}{\partial \sigma_i^2} \\
&= \sum_{n=1}^{N}\sum_{i=1}^{d}\mathbb{1}(y^{(n)}=k)(2\pi-\frac{1}{2\sigma_i^4}(x_i^{(n)}-\mu_{ki})^2)
\end{aligned}
$$

### 1.4 Estimation

$$
\begin{aligned}
&\frac{\partial \ell}{\partial \mu_{ki}} = 0 \\
&\Rightarrow \mu_{ki} = \frac{\sum_{n=1}^{N}\sum_{i=1}^{d}\mathbb{1}(y^{(n)}=k)x_i^{(n)}}{\sum_{n=1}^{N}\sum_{i=1}^{d}\mathbb{1}(y^{(n)}=k)} \\
&\frac{\partial \ell}{\partial \sigma_i^2} = 0 \\
&\Rightarrow \sigma_i = \sqrt[4]{\frac{\sum_{n=1}^{N}\sum_{i=1}^{d}\mathbb{1}(y^{(n)}=k)*(x_i^{(n)}-\mu_{ki})^2}{\sum_{n=1}^{N}\sum_{i=1}^{d}\mathbb{1}(y^{(n)}=k)*4\pi}}
\end{aligned}
$$