Introduction to Statistical Machine Learning
Homework 5

Yota Toyama

December 10, 2016

1.

$$
\begin{aligned}
p(y=1|\mathbf{x}) &= \frac{p(\mathbf{x}|y=1)p(y=1)}{p(\mathbf{x})} \\
&= \frac{p(y=1)}{p(\mathbf{x})} N(\mu_1, \Sigma_1) \\
&= \frac{p(y=1)}{p(\mathbf{x})} \prod_{j=1}^{d} N(\mu_{1,j}, \sigma_j) \\
&= \frac{p(y=1)}{p(\mathbf{x})} \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{1,j})^2}{2\sigma_j^2}\right) \\
&= \frac{p(y=1)\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{1,j})^2}{2\sigma_j^2}\right)}{p(y=0)\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{0,j})^2}{2\sigma_j^2}\right) + p(y=1)\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{1,j})^2}{2\sigma_j^2}\right)} \\
&= \frac{1}{1 + \frac{p(y=0)\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{0,j})^2}{2\sigma_j^2}\right)}{p(y=1)\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(x_j-\mu_{1,j})^2}{2\sigma_j^2}\right)}} \\
&= \frac{1}{1 + \frac{p(y=0)}{p(y=1)} \prod_{j=1}^{d} \exp\left((x_j-\mu_{1,j})^2 - (x_j-\mu_{0,j})^2\right)} \\
&= \frac{1}{1 + \exp\left(\log\frac{p(y=0)}{p(y=1)} + \sum_{j=1}^{d}(\mu_{1,j}^2 - \mu_{0,j}^2) + 2\sum_{j=1}^{d}(\mu_{0,j}-\mu_{1,j})x_j\right)}
\end{aligned}
$$

$\therefore$ The posterior $p(y=1|x)$ resulting from Gaussian generative model has the same form as the posterior in logistic regression model with the bias and weight below.

$$
\begin{cases}
w_0 = \sum_{j=1}^{d}(\mu_{1,j}^2 - \mu_{0,j}^2) \\
\mathbf{w} = 2(\mu_0 - \mu_1)
\end{cases}
$$

2. Yes, they produce the same classifier because it's a convex optimization problem.
3. Please, see a Jupyter notebook file submitted together.