

Homework

November 30, 2016

1 Genrative Model

1.1 Problem 1

In Gaussian generative model, we have:

$$\begin{aligned} p(y) &= \phi^y (1 - \phi)^{1-y} \\ p(x|y=0) &= \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu_0)^T \Sigma^{-1} (x - \mu_0)\right) \\ p(x|y=1) &= \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x - \mu_1)^T \Sigma^{-1} (x - \mu_1)\right) \end{aligned} \quad (1)$$

$$p(y=c|x) = \frac{p(x|y=c)p(y=c)}{\sum_c p(x|y=c)p(y=c)} \quad (2)$$

Suppose $c=0$, then we have

$$\begin{aligned} p(y=0|x) &= \frac{(1-\phi)\exp(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0))}{(1-\phi)\exp(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0)) + \phi \exp(-\frac{1}{2}(x-\mu_0)^T \Sigma^{-1} (x-\mu_0))} \\ &= \frac{(1-\phi)\exp(-\frac{1}{2} \sum_i \sigma_i^2 (x_i - \mu_{0i})^2)}{(1-\phi)\exp(-\frac{1}{2} \sum_i \sigma_i^2 (x_i - \mu_{0i})^2) + \phi \exp(-\frac{1}{2} \sum_i \sigma_i^2 (x_i - \mu_{1i})^2)} \\ &= \frac{1}{1 + \exp(\log \frac{\phi}{1-\phi} + \frac{1}{2} \sum_i (-2x_i \mu_{0i} + \mu_{0i}^2 + 2x_i \mu_{1i} - \mu_{1i}^2))} \\ &= \frac{1}{1 + \exp(\sum_i (\mu_{1i} - \mu_{0i})x_i + \frac{1}{2} \sum_i (\mu_{0i}^2 - \mu_{1i}^2) + \log \frac{\phi}{1-\phi})} \end{aligned} \quad (3)$$

Thus, for $c = 0$, with $\boldsymbol{\omega} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$ and $\omega_0 = \frac{1}{2} \sum_i (\mu_{0i}^2 - \mu_{1i}^2) + \log \frac{\phi}{1-\phi}$, $p(y = 0|x)$ can be represented as $p(y = 0|x) = \frac{1}{1+\exp(\boldsymbol{\omega}\mathbf{x}+\omega_0)}$

In a similar manner, we can derive $p(y = 1|x) = \frac{1}{1+\exp(\boldsymbol{\omega}\mathbf{x}+\omega_0)}$ while $\boldsymbol{\omega} = \boldsymbol{\mu}_0 - \boldsymbol{\mu}_1$ and $\omega_0 = \frac{1}{2} \sum_i (\mu_{1i}^2 - \mu_{0i}^2) + \log \frac{1-\phi}{\phi}$.

Therefore, $p(y = c|\mathbf{x})$ has the equivalent form with posterior of logistic regression.

1.2 Problem 2

The two classifiers will not produce the same classification boundary. The Gaussian model makes a stronger assumption: the data is of Gaussian distribution in each class(i.e. $(\mathbf{x}|y) \sim N(\boldsymbol{\mu}, \sigma)$). This assumption does not exist in logistic regression. We can assume other distributions of $(\mathbf{x}|y)$ while achieving the same form of $p(y|\mathbf{x})$ as logistic regression. Therefore, the generative Gaussian model has more parameters than logistic regression. If the dataset is large enough, Gaussian model will achieve better result. But logistic regression can give more general and robust results and is less likely to overfit given that it has less parameters.

2 Multilayer Neural Network