模式识别与机器学习作业

BraveY

2020 年 1 月 17 日

1 第三章有监督学习方法

1.1 题目1

Question: please show that the relationship between a discriminative classifier (say logistic regression) and the above specific class of Gaussian naive Bayes classifiers is precisely the form used by logistic regression.

1.1.1 解

依题意有如下证明：

$$\begin{aligned} p(Y=1|X) &= \frac{P(Y=1)P(X|Y=1)}{P(Y=1)P(X|Y=1)+P(Y=0)P(X|Y=0)} \\ &= \frac{1}{1+exp(ln\frac{P(Y=0)P(X|Y=0)}{P(Y=1)P(X|Y=1)})} \end{aligned} \tag{1}$$

因为有条件独立的假设,可以将连乘变为连加，并将高斯分布的公式带入有：

$$
\begin{aligned}
P(Y=1|X) &= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln\frac{P(x_i|Y=0)}{P(x_i|Y=1)})} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln\frac{\frac{1}{\sqrt{2\pi\sigma_i^2}}exp(\frac{-(x_i-\mu_{i0})^2}{2\sigma_i^2})}{\frac{1}{\sqrt{2\pi\sigma_i^2}}exp(\frac{-(x_i-\mu_{i1})^2}{2\sigma_i^2})})} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln\ exp(\frac{(x_i-\mu_{i1})^2-(x_i-\mu_{i0})^2}{2\sigma_i^2}))} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D(\frac{2(\mu_{i1}+\mu_{i0})x_i+\mu_{i1}^2-\mu_{i0}^2}{2\sigma_i^2}))} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D(\frac{(\mu_{i1}+\mu_{i0})x_i}{\sigma_i}+\frac{\mu_{i1}^2-\mu_{i0}^2}{2\sigma_i^2}))} \\
&= \frac{1}{1+exp(w_0+\sum_i^D w_ix_i)}
\end{aligned}
\tag{2}
$$

其中

$$
\begin{aligned}
w_0 &= ln\frac{1-\pi}{\pi}+\sum_i^D\frac{\mu_{i1}^2-\mu_{i0}^2}{2\sigma_i^2} \\
w_i &= \frac{(\mu_{i1}+\mu_{i0})x_i}{\sigma_i}
\end{aligned}
\tag{3}
$$

最后的形式就是逻辑回归的形式，因此可以从高斯朴素贝叶斯推出逻辑回归的形式。

1.2 题目2

Question: is the new form of $P(y|x)$ implied by this more general Gaussian naive Bayes classifier still the form used by logistic regression? Derive the new form of $P(y|x)$ to prove your answer.

1.2.1 解

根据题1的推导有：

$$
\begin{aligned}
P(Y=1|X) &= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln\frac{P(x_i|Y=0)}{P(x_i|Y=1)})} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln\frac{\frac{1}{\sqrt{2\pi\sigma_{i0}^2}}exp(\frac{-(x_i-\mu_{i0})^2}{2\sigma_{i0}^2})}{\frac{1}{\sqrt{2\pi\sigma_{i1}^2}}exp(\frac{-(x_i-\mu_{i1})^2}{2\sigma_{i1}^2})})} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+\sum_i^D ln(\frac{\sigma_{i1}}{\sigma_{i0}}\ exp(\frac{(x_i-\mu_{i1})^2}{2\sigma_{i1}^2}-\frac{(x_i-\mu_{i0})^2}{2\sigma_{i0}^2}))}
\end{aligned}
\tag{4}
$$

因为x_i包含二项式，不是线性的关系所以不再是逻辑回归的形式。

1.3 题目3

Question: is the form of $P(y|x)$ implied by such not-so-naive Gaussian Bayes classifiers still the form used by logistic regression? Derive the form of $P(y|x)$ to prove your answer.

1.3.1 解

同题1前面的推导一样有：

$$
\begin{aligned}
p(Y=1|X) &= \frac{P(Y=1)P(X|Y=1)}{P(Y=1)P(X|Y=1)+P(Y=0)P(X|Y=0)} \\
&= \frac{1}{1+exp(ln\frac{P(Y=0)P(X|Y=0)}{P(Y=1)P(X|Y=1)})} \\
&= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+ln\frac{P(X|Y=0)}{P(X|Y=1)})}
\end{aligned}
\tag{5}
$$

将$P(X|Y)$的公式带入推导有：

$$\begin{aligned}\frac{P(X|Y=0)}{P(X|Y=1)} = exp(&\frac{2x_1(\sigma_2^2(\mu_{10}-\mu_{11})+\rho\sigma_1\sigma_2(\mu_{21}-\mu_{20}))+2x_2(\sigma_1^2(\mu_{20}-\mu_{21})+\rho\sigma_1\sigma_2(\mu_{11}-\mu_{10}))}{2(1-\rho)^2\sigma_i^2\sigma_2^2} \\ &+\frac{(\sigma_2\mu_{11}-\sigma_1\mu_{21})^2-(\sigma_2\mu_{10}-\sigma_1\mu_{20})^2+2(1-\rho)\sigma_1\sigma_2(\mu_{21}\mu_{11}-\mu_{20}\mu_{10})}{2(1-\rho)^2\sigma_i^2\sigma_2^2})\end{aligned} \tag{6}$$

带入上式可以写成：

$$\begin{aligned} p(Y=1|X) &= \frac{1}{1+exp(ln\frac{1-\pi}{\pi}+ln\frac{P(X|Y=0)}{P(X|Y=1)})} \\ &= \frac{1}{1+exp(w_1x_1+w_2x_2+w_0)}\end{aligned} \tag{7}$$

其中

$$\begin{aligned} w_1 &= \frac{2(\sigma_2^2(\mu_{10}-\mu_{11})+\rho\sigma_1\sigma_2(\mu_{21}-\mu_{20}))}{2(1-\rho)^2\sigma_i^2\sigma_2^2} \\ w_2 &= \frac{2(\sigma_1^2(\mu_{20}-\mu_{21})+\rho\sigma_1\sigma_2(\mu_{11}-\mu_{10}))}{2(1-\rho)^2\sigma_i^2\sigma_2^2} \\ w_0 &= \frac{(\sigma_2\mu_{11}-\sigma_1\mu_{21})^2-(\sigma_2\mu_{10}-\sigma_1\mu_{20})^2+2(1-\rho)\sigma_1\sigma_2(\mu_{21}\mu_{11}-\mu_{20}\mu_{10})}{2(1-\rho)^2\sigma_i^2\sigma_2^2}+ln\frac{1-\pi}{\pi}\end{aligned} \tag{8}$$

因此仍然满足逻辑回归的形式。