MACHINE LEARNING

Homework Week 6

October 13, 2021

Nguyen Son Tung

---

1. Logistic Regression

Answer

Consider first of all the case of two classes C_1 and $C2$. The posterior probability for class C1 can be written as:

$$p(C_1|x) = \frac{p(x|C_1)p(C_1)}{p(x|C_1)p(C_1) + p(x|C_2)p(C_2)} = \frac{1}{1+e^{-a}} = \sigma(a)$$

where we have defined

$$a = \log \frac{p(x|C_1)p(C_1)}{p(x|C_2)p(C_2)}$$

and $\sigma(a)$ is the logistic sigmoid function defined by

$$\begin{aligned}
\sigma(a) &= \frac{1}{1+e^{-a}} \\
\frac{\partial \sigma(a)}{\partial a} &= \frac{e^{-a}}{(1+e^{-a})^2} \\
&= \frac{1}{(1+e^{-a})} \cdot \frac{e^{-a}}{(1+e^{-a})} \\
&= \frac{1}{(1+e^{-a})} \cdot \frac{1+e^{-a}-1}{(1+e^{-a})} \\
&= \frac{1}{(1+e^{-a})} \cdot \left(\frac{1+e^{-a}}{1+e^{-a}} - \frac{1}{1+e^{-a}}\right) \\
&= \sigma(a) \cdot (1-\sigma(a))
\end{aligned}$$

The model logistic regression is defined as:

$$p(C_1|\phi) = y(\phi) = \sigma(w^T\phi)$$

$$p(C_2|\phi) = 1 - p(C_1|\phi)$$

For a data set ϕ_n, t_n, where $t_n \in \{0,1\}$ and $\phi_n = \phi(x_n)$, with n = 1,...,N, the likelihood function can be written as

$$p(t|w) = \prod_{n=1}^{N} y_n^{t_n}(1-y_n)^{1-t_n}$$

where t = $(t_1, ..., t_N)^T$ and $y_n = p(C_1|\phi_n)$ We can define an error function by taking the negative logarithm of the likelihood, which gives the cross_entropy error function in the form

$$L = -\log p(t|w) = -\sum_{n=1}^{N}\{t_n \log y_n + (1-t_n)\log(1-y_n)\}$$

with $y_n = \sigma(a_n)$ and $a_n = w^T\phi_n$
To maximize L we need to calculate the derivative and let it = 0. Then,

$$L = t\log y + (1-t)\log(1-y)$$

with

$$y = \sigma(z)$$

and

$$z = w_0 + w_1\phi_1 + w_2\phi_2 + \cdots + w_n\phi_n$$

Using the Chain rule, we have:

$$\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial y}\cdot\frac{\partial y}{\partial z}\cdot\frac{\partial z}{\partial w_i}$$

Calculate:

$$\frac{\partial L}{\partial y} = \frac{t}{y} - \frac{1-t}{1-y} = \frac{y-t}{y(1-y)} \tag{1}$$

$$\frac{\partial y}{\partial z} = y(1-y) \tag{2}$$

$$\frac{\partial z}{\partial w_i} = \phi_i \tag{3}$$

According to the Chain rule:

$$\frac{\partial L}{\partial w_i} = \frac{y-t}{y(1-y)}\cdot y(1-y)\cdot\phi_i$$

$$= (y - t)\phi_i$$

or

$$\nabla L = \sum_{n=1}^{N} (y_n - t_n)\phi_n$$

2. Find $f(x)$ that $f'(x) = f(x)(1 - f(x))$

Answer

We have:

$$f'(x) = f(x)(1 - f(x))$$

$$\Leftrightarrow \frac{\partial f(x)}{\partial x} = f(x)(1 - f(x))$$

$$\Leftrightarrow \partial x = \frac{\partial f(x)}{f(x)(1 - f(x))}$$

Integrating both sides of the equation, we get:

$$\int \partial x = \int \frac{\partial f(x)}{f(x)(1 - f(x))}$$

$$\Leftrightarrow x = \int \left(\frac{1}{f(x)} + \frac{1}{1 - f(x)} \right) \partial f(x)$$

$$\Leftrightarrow x = \int \frac{1}{f(x)} \partial f(x) + \frac{1}{1 - f(x)} \partial f(x)$$

$$\Leftrightarrow x = \log f(x) + \log(1 - f(x))$$

$$\Leftrightarrow x = \log \left(\frac{f(x)}{1 - f(x)} \right)$$

$$\Leftrightarrow e^x = \frac{f(x)}{1 - f(x)}$$

$$\Leftrightarrow f(x) = e^x - e^x f(x)$$

$$\Leftrightarrow f(x) + e^x f(x) = e^x$$

$$\Leftrightarrow f(x)(1 + e^x) = e^x$$

$$\Leftrightarrow f(x) = \frac{e^x}{1 + e^x}$$