Machine Learning Exercise sheet 5

Linear Classification

Anastasia Stamatouli : 03710902

17/10/2019

- Linear Classification

Problem (1):

a) exponential distribution

$$p(y=c \mid x) \propto p(x \mid y=c) \cdot p(y=c) \Leftrightarrow$$

$$p(y=0 \mid x) \propto \frac{\lambda_0}{2} e^{-\lambda_0 x}$$ (which is an exponential distribution)

b) x is classified as 1 when:

$$p(y=1 \mid x) > p(y=0 \mid x)$$

$$p(y=1 \mid x) = \frac{\frac{1}{2}\lambda_1 e^{-\lambda_1 x}}{\frac{1}{2}\lambda_1 e^{-\lambda_1 x} + \frac{1}{2}\lambda_0 e^{-\lambda_0 x}}$$

$$p(y=0 \mid x) = \frac{\frac{1}{2}\lambda_0 e^{-\lambda_0 x}}{\frac{1}{2}\lambda_0 e^{-\lambda_0 x} + \frac{1}{2}\lambda_1 e^{-\lambda_1 x}}$$

$$\frac{\frac{1}{2}\lambda_1 e^{-\lambda_1 x}}{\frac{1}{2}\lambda_1 e^{-\lambda_1 x} + \frac{1}{2}\lambda_0 e^{-\lambda_0 x}} > \frac{\frac{1}{2}\lambda_0 e^{-\lambda_0 x}}{\frac{1}{2}\lambda_0 e^{-\lambda_0 x} + \frac{1}{2}\lambda_1 e^{-\lambda_1 x}} \Leftrightarrow$$

$$(\lambda_1 e^{-\lambda_1 x})^2 > (\lambda_0 e^{-\lambda_0 x})^2 \Leftrightarrow$$

$$\left(\frac{\lambda_1}{\lambda_0}\right)^2 > \frac{e^{-2\lambda_0 x}}{e^{-2\lambda_1 x}} \Leftrightarrow$$

$$\ln\left(\frac{\lambda_1}{\lambda_0}\right)^2 > \ln\left(e^{-2\lambda_0 x - 2\lambda_1 x}\right) \Leftrightarrow$$

$$\ln\left(\frac{\lambda_1}{\lambda_0}\right)^2 > -2(\lambda_0+\lambda_1)x \quad \Leftrightarrow \quad \lambda_0+\lambda_1>0$$ since λ_0, λ_1 parameter of exp distribution

$$x < \frac{-\ln\left(\frac{\lambda_1}{\lambda_0}\right)}{\lambda_0-\lambda_1}$$

Problem ②:

sigmoid function: $\dfrac{1}{1+e^{-wx}}$

If we plot the sigmoid function for increasing values of w, we observe that the curve gets steeper as w increases. Steeper curve means that the model is almost sure about the class. (almost 0 probability or almost 1), which means overfitting.

b) To prevent overfitting we want the weights to be small. To achieve this, instead of maximum conditional likelihood estimation we can consider maximum conditional a posterior where we assume Gaussian prior for the weight vector.

Problem 3:

softmax : $\sigma(x)_i = \dfrac{\exp(x_i)}{\sum_{k=1}^{K} \exp(x_k)}$

$K = 2$

$$\sigma(x)_i = \frac{e^{x_0}}{e^{x_0}+e^{x_1}} = \frac{\frac{e^{x_0}}{e^{x_0}}}{\frac{e^{x_0}+e^{x_1}}{e^{x_0}}} = \frac{1}{1+\frac{e^{x_1}}{e^{x_0}}} =$$

$$\frac{1}{1+e^{x_1-x_0}} \quad (1)$$

. Set $x_1 - x_0 = z$

$$(1) = \frac{1}{1+e^{z}} = \text{sigmoid}(z) = \text{sigmoid}(x_1 - x_0)$$

Problem 4:

	x_1	x_2
We notice that circles have	negative positive	negative positive
crosses have	positive	negative

so if the transformation is $\varphi(x_1 x_2, x_2)$ one group (circles) will have positive x while the other group (crosses) negative.