BRUNO BORGES DE SOUZA

LISTA 4

Exercício 17.7

Na regressão logística, para cada ponto nos dados temos o vetor $x_i$ e uma classe observada $y_i$. A probabilidade P, se $y_i = 1$ ou $1-P$ se $y_i = 0$. A verossimilhança é dada então, por:

$$L = \prod_{i=1}^{m} P(x_i)^{y_i} (1 - P(x_i))^{1-y_i}$$

Como $P(c=1|x) = \sigma(w^T x + b)$, sendo $D \doteq \{(x^n, c^n), n = 1, \dots, N\}$ $c^n \in \{0,1\}$, temos:

$$L = \prod_{i=1}^{n} \sigma(w^T x^n + b)^{c^i} (1 - \sigma(w^T x^n + b))^{1-c^i}$$

$$\log L = \sum_i c^i \log \sigma(w^T x^n + b) + (1-c^i) \log(1 - \sigma(w^T x^n + b))$$

Como a derivada de $\log_e \sigma(w^T x^n + b) = \dfrac{\sigma'}{\sigma(w^T x^n + b)}$ (com relação a w), e observando que:

$$\sigma(x) = \frac{1}{1+e^{-\lambda x}}, \therefore \sigma'(x) = \frac{-\lambda e^{-\lambda x}}{(1+e^{-\lambda x})^2} = \sigma(x)^2(-\lambda e^{-\lambda x}),$$

$$\lambda e^{-\lambda x} = \frac{\lambda(1-\sigma)}{\sigma} \Rightarrow \sigma' = \sigma(1-\sigma), \text{ temos}$$

$$\nabla_w L = \sum_i c^i (1 - \sigma(w^T x^i + b)) x^i - (1-c^i) \sigma(w^T x^i + b) x^i$$

Com $i = n$ e simplificando a expressão acima:

$$\nabla_w L = \sum_{n=1}^{N} (c^n - \sigma(w^T x^n + b)) x^n \quad \blacksquare$$