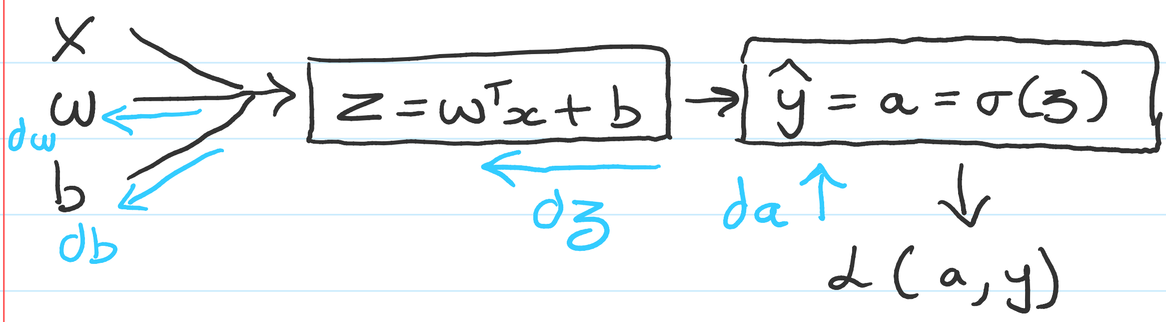
Logistic Regression with Gradient descent

30 December 2018 02:33

$$z = \omega^T x + b$$

$$\hat{y} = a = \sigma(z)$$

$$\mathcal{L}(a,y) = -\left(y \log(a) + (1-y)\log(1-a)\right)$$



$$da = d\mathcal{L}(a,y)/da$$

$$= d\left(-y \log a - (1-y)\log(1-a)\right)/da$$

$$= -y \frac{d}{da} \log a + (y-1) \frac{d}{da} \log(1-a)$$

$$= -y/a + (y-1) \frac{1 \quad \frac{d}{da}(1-a)}{(1-a)}$$

$$= -y/a + \frac{y-1 \quad (0-1)}{(1-a)}$$

$$= -y/a + \frac{(1-y)}{(1-a)}$$

$$= \frac{-y(1-a) + a(1-y)}{a(1-a)}$$

$$= -y \cancel{+ya} + a \cancel{-ay} \,/\, a(1-a)$$

$$= (a-y) \,/\, a(1-a)$$

$$dz = \partial \mathcal{L}(a,y) / \partial z$$

$$\left[\frac{1}{U(x)}\right]' = \frac{U'(x)}{U(x)^2}$$

$$= \frac{\partial \mathcal{L}(a,y)}{\partial a} \times \partial a / \partial z$$

$$x = 1 + e^{-z}$$

$$\partial a / \partial z = \partial(\sigma(z)) / \partial z$$

$$= \partial / \partial z \left(\frac{1}{1+e^{-z}}\right)$$

$$= -\frac{\partial / \partial z \, (1+e^{-z})}{(1+e^{-z})^2}$$

$$= -\frac{\partial / \partial z \, (1) + \partial / \partial z \, (e^{-z})}{(1+e^{-z})^2}$$

$$= -\frac{0 + e^{-z} \, \partial / \partial z \, (-z)}{(1+e^{-z})^2}$$

$$= e^{-z} / (1+e^{-z})^2$$

$$= \frac{1}{(1+e^{-z})} \frac{(e^{-z}+1-1)}{1+e^{-z}}$$

$$= \frac{1}{1+e^{-z}} \left(\left(\frac{1+e^{-z}}{1+e^{-z}}\right) - \frac{1}{1+e^{-z}} \right)$$

$$= \sigma(z)(1-\sigma(z))$$

$$= a(1-a)$$

$$dz = (a-y) / \cancel{a(1-a)} \times \cancel{a(1-a)}$$

$$= a - y$$

$$= a - y$$

$$d\omega = \frac{\partial \mathcal{L}(a,y)}{\partial \omega}$$

$$= \frac{\partial \mathcal{L}(a,y)}{\partial z} \times \frac{\partial z}{\partial \omega}$$

$$\partial z / \partial \omega = \frac{\partial (\omega^T x + b)}{\partial \omega}$$

$$= x$$

$$d\omega = x \; \partial(\mathcal{L}(a,y)) / \partial z$$

$$= x \; dz$$

$$db = \partial \mathcal{L}(a,y) / \partial b$$

$$= \frac{\partial \mathcal{L}(a,y)}{\partial z} \times \partial z / \partial b$$

$$= \partial(\omega^T x + b) / \partial b$$

$$= 1$$

$$db = \partial(\mathcal{L}(a,y)) / \partial z$$

$$= dz$$

Update weights & bias

$$\omega := \omega - \alpha \, d\omega$$

$$b := b - \alpha \, db$$