- 针对m个样本的梯度下降

针对m个样本的梯度下降

我们仍然假设样本只有两个参数

首先列出损失函数

$$J(w,b) = -\frac{1}{m}\sum_{i=1}^{m} L(a^{(i)}, y)$$

其次可以得到

$$a^{(i)} = \hat{y}^{(i)} = \sigma(z^{(i)}) = \sigma(w^T\vec{x}^{(i)} + b)$$

所以可以得出

$$\frac{\partial J(w,b)}{\partial w_1} = \frac{1}{m}\sum_{i=1}^{m}\frac{\partial L(a^{(i)}, y)}{\partial w_1^{(i)}} = \frac{1}{m}\sum_{i=1}^{m} x_1^{(i)} dz^{(i)}$$

$$\frac{\partial J(w,b)}{\partial w_2} = \frac{1}{m}\sum_{i=1}^{m}\frac{\partial L(a^{(i)}, y)}{\partial w_2^{(i)}} = \frac{1}{m}\sum_{i=1}^{m} x_2^{(i)} dz^{(i)}$$

$$\frac{\partial J(w,b)}{\partial b} = \frac{1}{m}\sum_{i=1}^{m}\frac{\partial L(a^{(i)}, y)}{\partial b^{(i)}} = \frac{1}{m}\sum_{i=1}^{m} dz^{(i)}$$

所以我们现在对参数进行初始化

$$J = 0, dw_1 = 0, dw_2 = 0, db = 0$$

现在,使用**for循环**从1到m遍历训练集,同时计算相应的每个训练样本的导数,分别求和遍历完成后再将得到的和分别除以m,就计算出了损失函数$J(w,b)$以及对各个参数的偏导数

使用得到的导数对各个参数进行一次迭代,得到新的参数,再使用新的参数重复上述过程,进行若干次直到损失函数收敛