3.10

$$\frac{dL}{dz^{[2]}} = \frac{dL}{da^{[2]}} \frac{da^{[2]}}{dz^{[2]}} = a^{[2]} - y$$

$$\frac{dL}{dw^{[2]}} = \frac{dL}{dz^{[2]}} \cdot \frac{dz^{[2]}}{dw^{[2]}} = (a^{[2]} - y) \cdot a^{[1]T}$$

$$\frac{dL}{db} = \frac{dL}{dz^{[2]}} \cdot 1 = dz^{[2]}$$

$$\frac{dL}{dz^{[1]}} = \frac{dL}{dz^{[2]}} \cdot \frac{dz^{[2]}}{dz^{[1]}}$$

$$= \frac{dL}{dz^{[2]}} \cdot \frac{dz^{[2]}}{da^{[1]}} \cdot \frac{da^{[1]}}{dz^{[1]}}$$

$$= \frac{dL}{dz^{[2]}} \cdot w^{[2]}.$$

$$= w^{[2]T} dz^{[2]} \cdot g^{[1]\prime}(z^{[1]})$$

$$dw^{[1]} = dz^{[1]} x^T$$

$$db^{[1]} = dz^{[1]}$$

2019 09 24.

为什么无偏估计的方差是除以 $n-1$?

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$E(XY) = EX \cdot EY$$

$$Var(\bar{X}) = \frac{1}{n} Var(X) \quad ?$$

$$\sigma(\bar{x})^2 = \frac{1}{n} \sigma(x)^2$$

二阶中心距：

$$\boxed{Var(x) = E(x^2) - (E(x))^2}$$

$$E(X^2) = Var(X) + (E(x))^2 = \sigma^2 + \mu^2$$

$$E(\bar{X}^2) = Var(\bar{x}) + (E(\bar{x}))^2 = \frac{1}{n}\sigma^2 + \mu^2$$

无偏估计：对变量θ的估计是$\hat{\theta}$，如果 $E(\hat{\theta}) = E(\theta)$，则称$\hat{\theta}$为θ的无偏估计。

$$E\left(\sum_{i=1}^{n} (x_i - \bar{x})^2\right)$$

$$= \sum_{i=1}^{n} E(x_i^2) - n \cdot E(\bar{x}^2)$$

$$= \sum_{i=1}^{n} (\sigma^2 + \mu^2) - n\left(\frac{1}{n}\sigma^2 + \mu^2\right)$$

$$= (n-1)\sigma^2$$