

17.03.2022 | neural networks

$$y_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = f(x_i), \quad \text{data } y_1, \dots, y_n, x_1, \dots, x_n$$

$$P(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(f(x_i) - y_i)^2}{2\sigma^2}}$$

$$\log P(y_1, \dots, y_n | x_1, \dots, x_n) = \sum_{i=1}^n -\log(\sqrt{2\pi}\sigma) -$$

$$- \frac{(f(x_i) - y_i)^2}{2\sigma^2}$$

$$\arg \max_{\vec{\beta}, \vec{L}} \log P(y_1, \dots, y_n | x_1, \dots, x_n)$$

$$= \arg \max_{\vec{\beta}, \vec{L}} \left(\sum_{i=1}^n -\log(\sqrt{2\pi}\sigma) - \right.$$

$$\left. - \frac{(f(x_i) - y_i)^2}{2\sigma^2} \right) =$$

$$= \arg \max_{\vec{\beta}, \vec{L}} - \sum_{i=1}^n (f(x_i) - y_i)^2$$

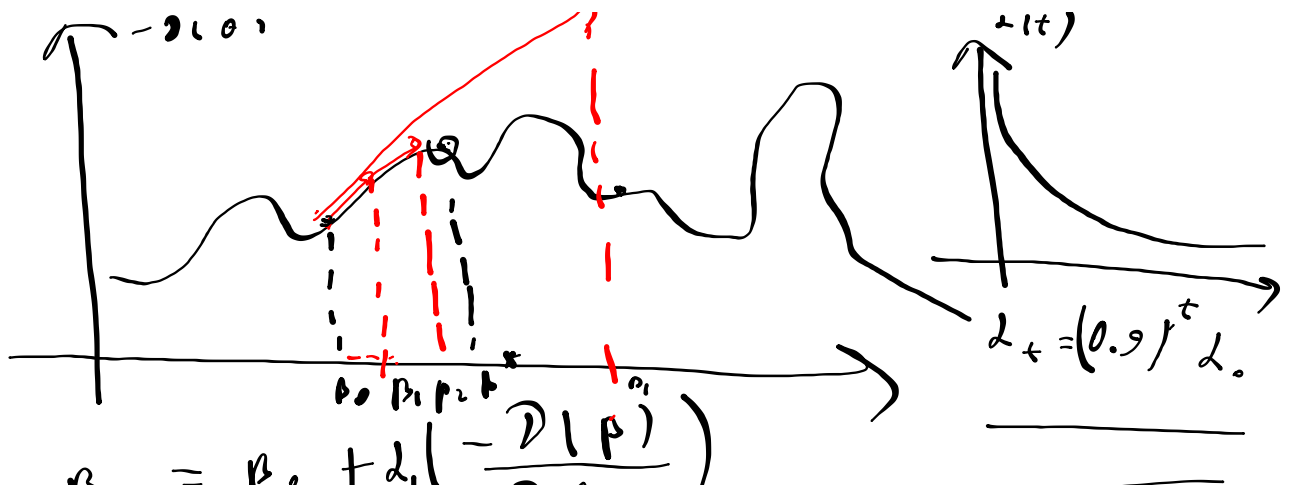
$y_i \overset{i.i.d.}{\sim} \text{Bernoulli}(p_i)$

$p_i = f(x_i)$. for the data $y_1, \dots, y_n, x_1, \dots, x_n$

$$P(y_1, \dots, y_n | x_1, \dots, x_n) = \prod_{i=1}^n y_i^{f(x_i)} (1 - y_i)^{1 - f(x_i)}$$

$$\log P(y_1, \dots, y_n | x_1, \dots, x_n) =$$

$$= \sum_{i=1}^n f(x_i) \log y_i + (1 - f(x_i)) \log (1 - y_i)$$



$$\beta_1 = \beta_0 + \lambda_1 \left(-\frac{\partial f(\beta)}{\partial \beta} \right)$$

$$\beta_2 = \beta_1 + \lambda_2 \left(-\frac{\partial f(\beta)}{\partial \beta} \right)$$

$$\ell(\beta) = -f(\beta)$$

$$\hat{\ell}(\beta) = \frac{\ell^{(0)}(\beta)}{0!} + \frac{\ell'(\beta)\Delta}{1!} + \frac{\ell''(\beta)\Delta^2}{2!} =$$

$$\frac{\partial \hat{\ell}}{\partial \Delta}(\beta) = \ell'(\beta) + \Delta \ell''(\beta) (=0)$$

$$\Rightarrow \Delta = -\frac{\ell'(\beta)}{\ell''(\beta)} = -\frac{\partial f(\beta)}{\partial^2 f(\beta)}$$