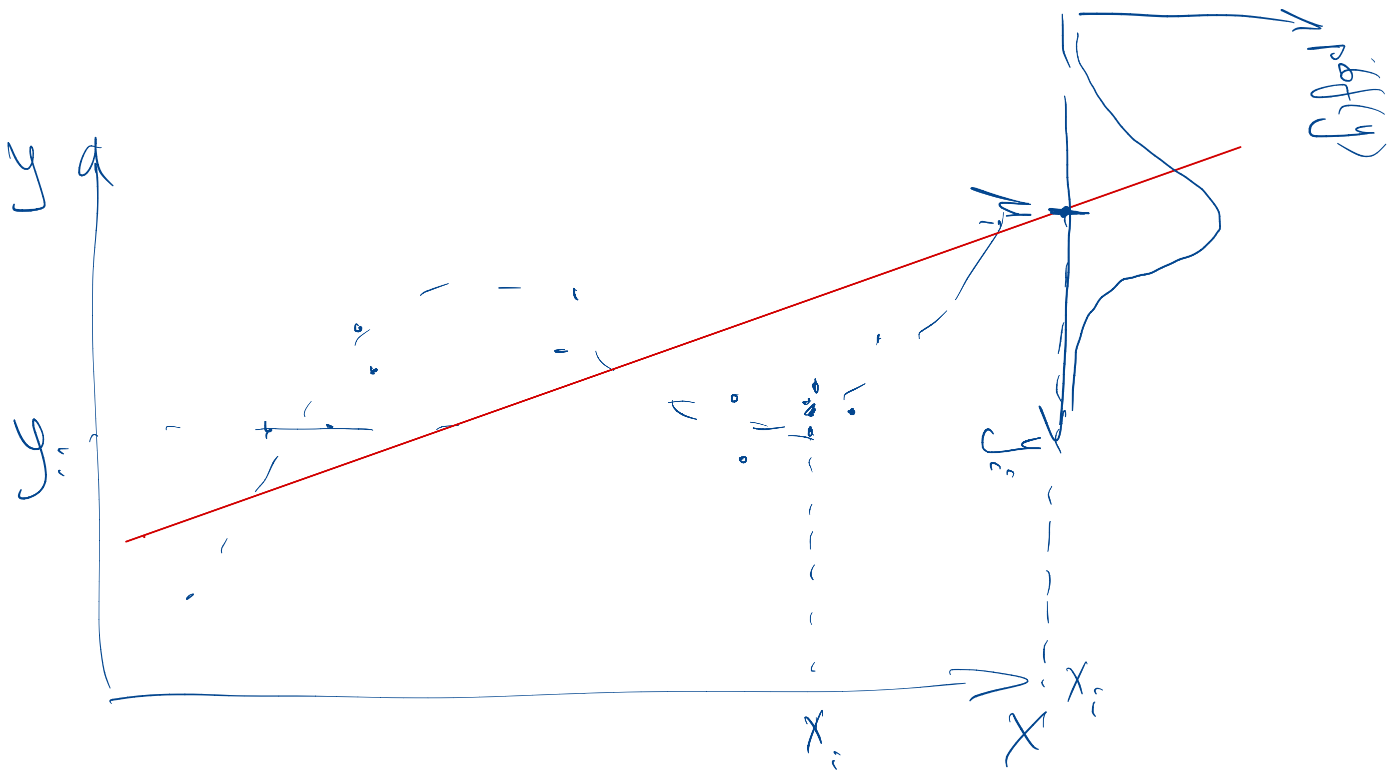
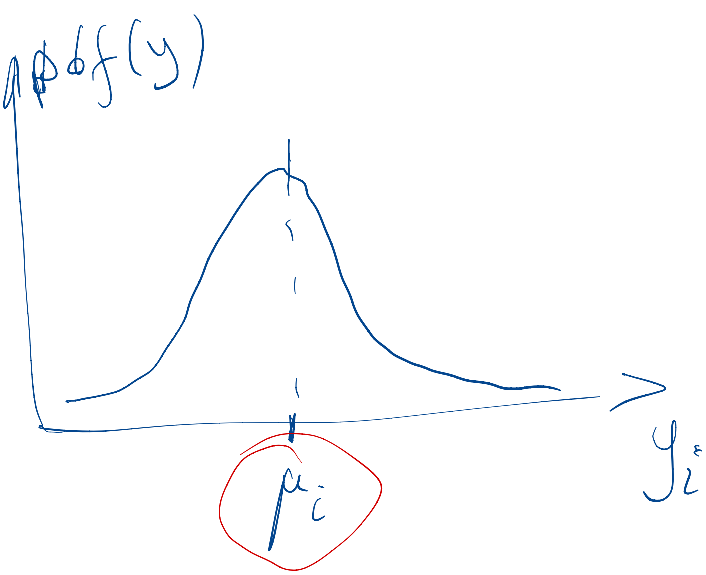
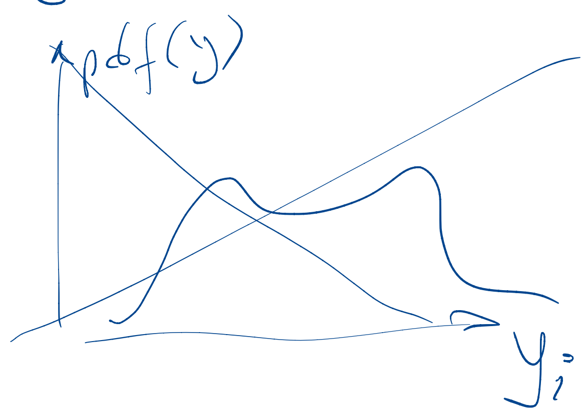
# Линейная регрессия




$$T = \{(x_i, y_i)\} \qquad y_i \sim P(y, x_i; \Theta)$$



$$y_i = \mathcal{N}(\mu_i, \sigma_i^2) \quad \mu_i = \theta^T x_i$$

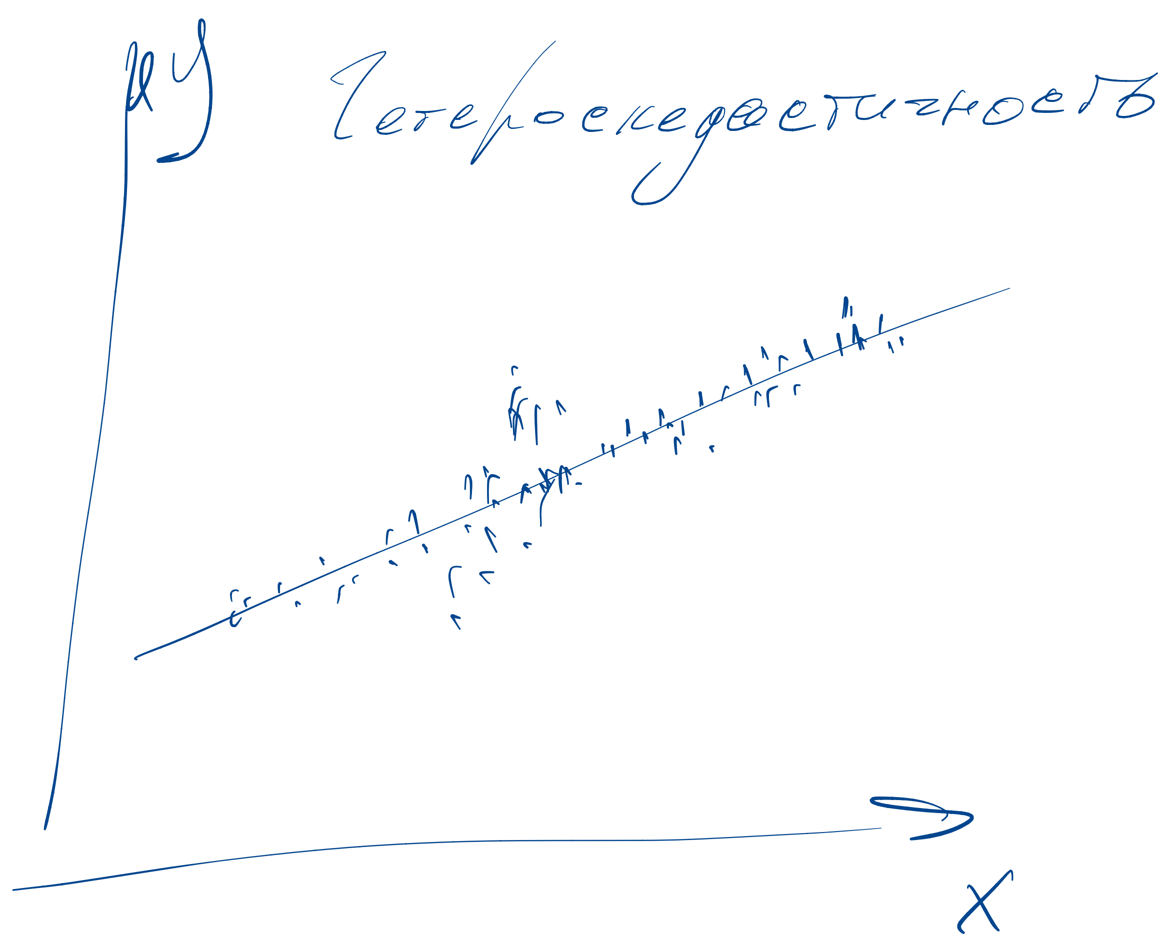
# LINE

N: $y_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$

L: $\mu_i = \theta^T x_i$

I: i.i.d.: independent; identically distributed

E: equivariance $\sigma_i^2 = \sigma^2$ - Гомоскедастичность

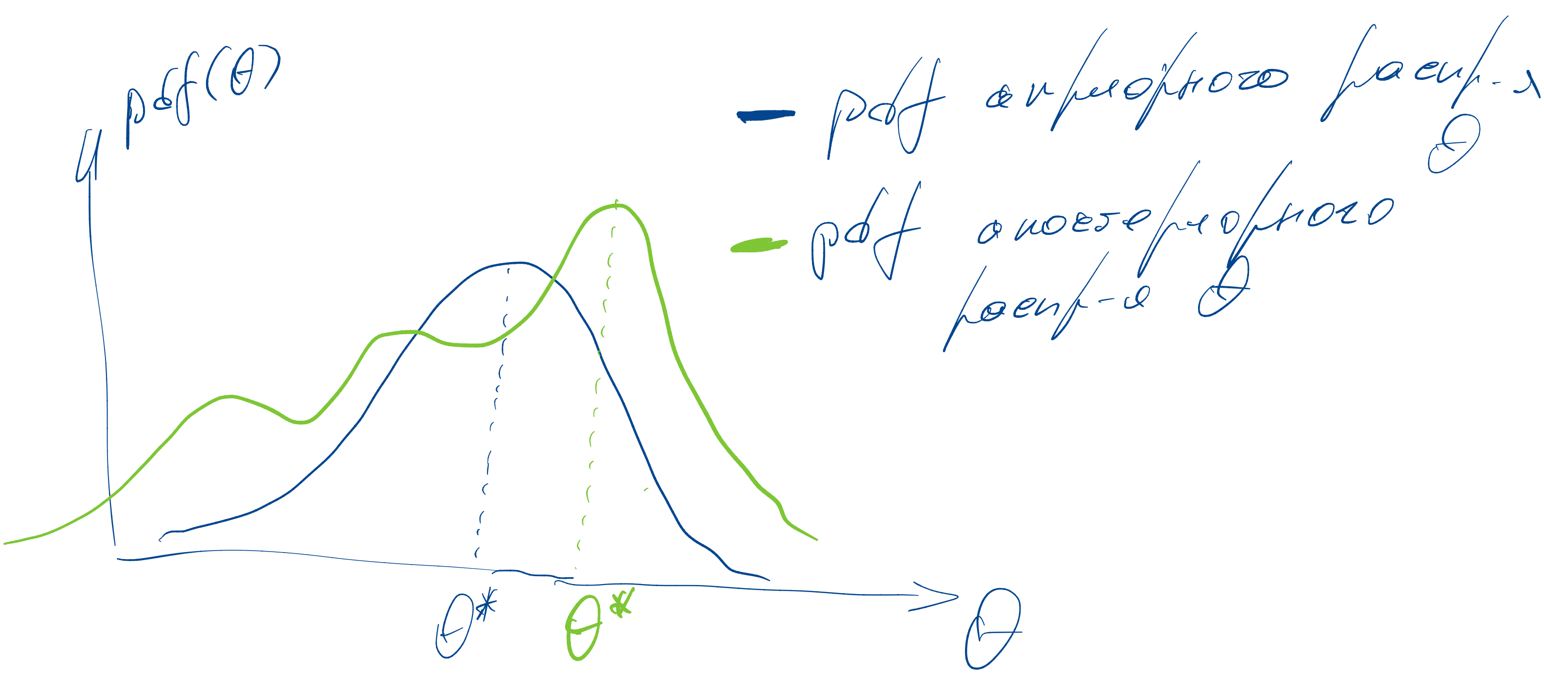


# Метод максимизации правдоподобия

## Формула Байеса

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{P(B|A)\,P(A)}{P(B)}$$

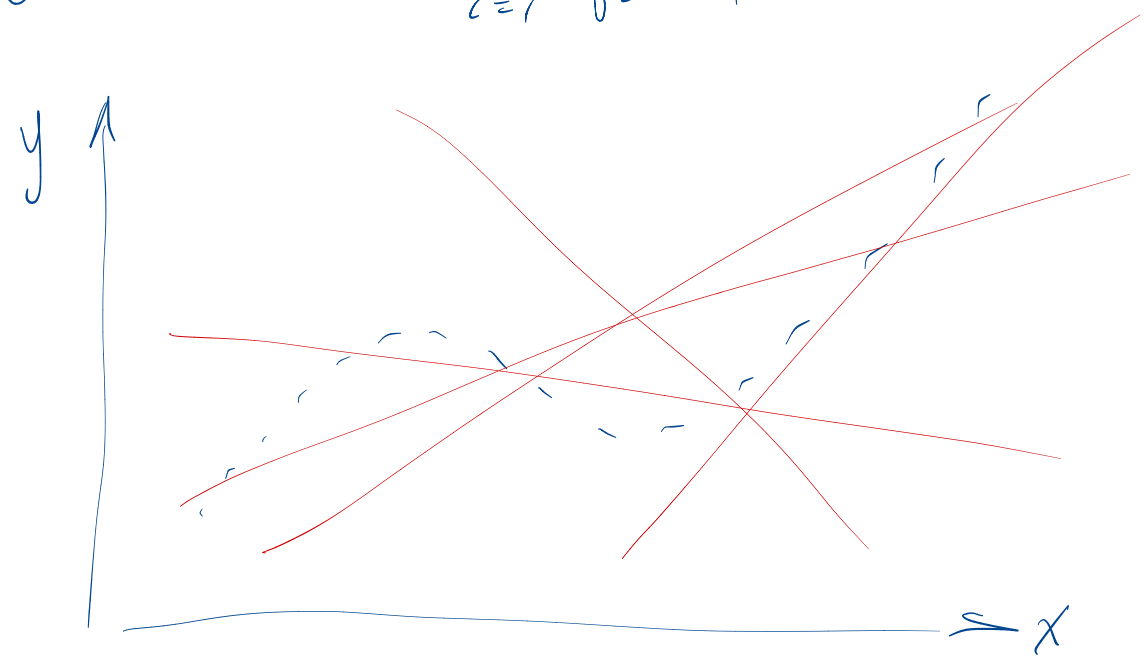
$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)\,P(\theta)}{P(\mathcal{D})}$$



$$P_i = \frac{1}{\sqrt{2\pi \sigma_i^2}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} = P((x_i, y_i) | \theta)$$

$$P(T|\theta) - ? = \prod_{i=1}^{N} P_i = \prod_{i=1}^{N} \frac{1}{\sqrt{\dots}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}}$$

$$\mathcal{L}(T, \theta) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi \sigma_i^2}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}}$$




$$\theta^* = \underset{\Theta}{\operatorname{argmax}} \mathcal{L}(T, \theta)$$

$$\nabla_\theta \mathcal{L} = 0 \qquad \nabla_\theta \ln \mathcal{L} = 0$$

$$\ln \mathcal{L} = \sum \ln \left( \frac{1}{\sqrt{2\pi \sigma_i^2}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} \right)$$

$$\theta^* = \underset{\Theta}{\arg\max} \left( \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} + \sum_{i=1}^{N} \ln e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} \right)$$

$$\underset{\Theta}{\arg\max}\, \mathcal{L} = \underset{\Theta}{\arg\max}\, \ln(\mathcal{L})$$

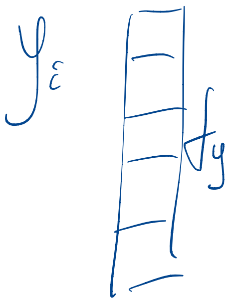
$$\theta^* = \underset{\Theta}{\arg\max} \left( \sum - \frac{(y_i-\mu_i)^2}{2\sigma_i^2} \right) =$$

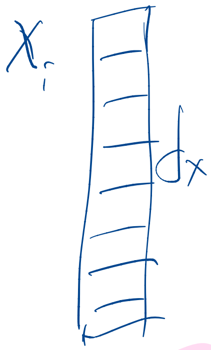
$$= \underset{\Theta}{\arg\max}\, \cancelto{1}{\frac{1}{2\sigma^2}} \sum \left(-(y_i-\mu_i)^2\right)$$

$$\theta^* = \underset{\Theta}{\arg\max} \left[ -\sum (y_i-\mu_i)^2 \right]$$

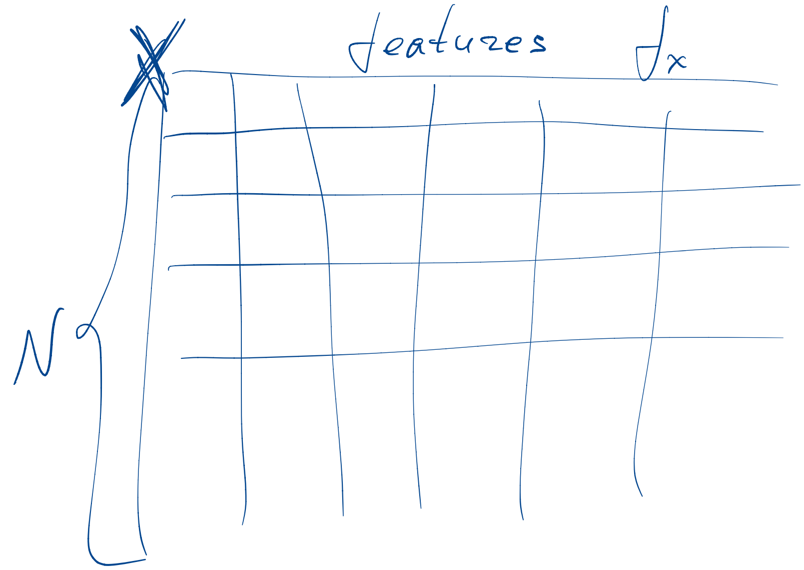
$$\theta^* = \underset{\Theta}{\arg\min} \sum (y_i-\mu_i)^2$$

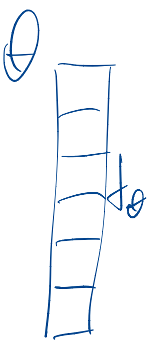
$$L(y, \theta) = \sum (y_i - \theta^T x_i)^2$$




$$\{x_i\} = X^T$$

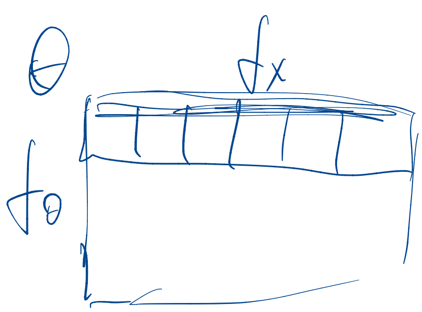


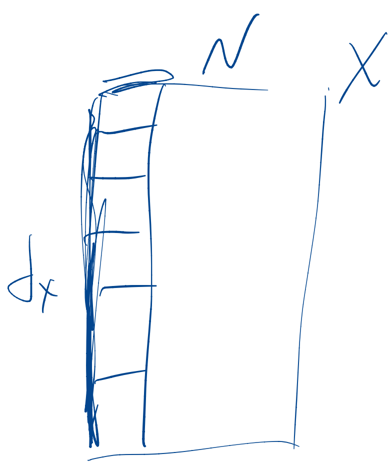




$$\theta x = y$$

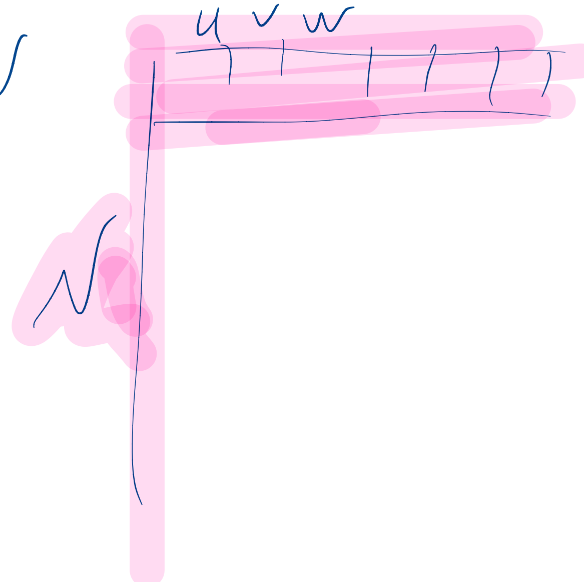
$$\underbrace{d_y \times N}_{Y} = \underbrace{d_\theta \times d_x}_{\theta} \odot \underbrace{d_x \times N}_{X}$$

$$\theta : d_y \times d_x$$



$$\mu = \theta^T X \qquad \mu : d_y \times N$$




$$L = (Y - \theta^T X)^T (Y - \theta^T X)$$

$$= (Y - X\theta)^T (Y - X\theta)$$

$$= Y^T Y - 2 Y^T X \theta + \theta^T X^T X \theta =$$

$$= Y^T Y - 2\theta^T X^T Y - \theta^T X^T X \theta$$

$$\nabla_\theta L = 0 - 2 X^T Y + 2 X^T X \theta$$

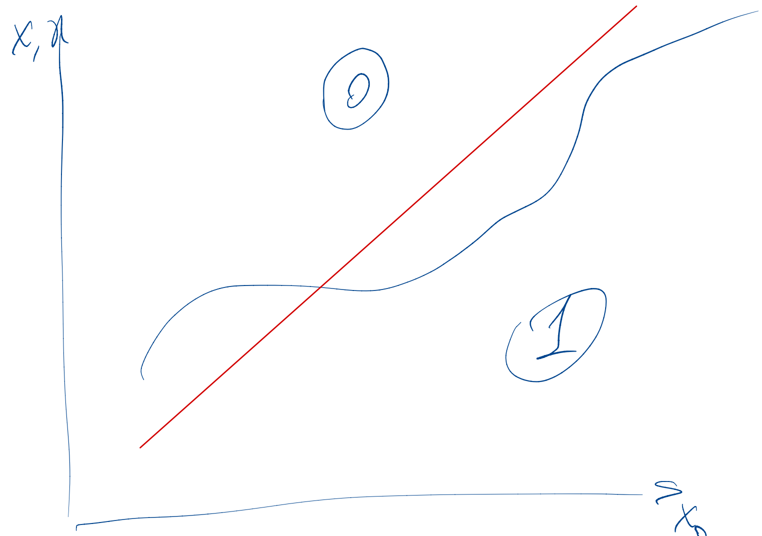
$$\nabla_\theta L = 0$$

$$\underline{2 X^T X \theta = 2 X^T Y}$$

$$\boxed{\theta^* = (X^T X)^{-1} X^T Y}$$

Задача классификации.
Модель логистической регрессии.

$y: 0 / 1$




$y_i \sim \mathcal{B}(p_i)$  $p_i = p_i(\theta, x_i)$

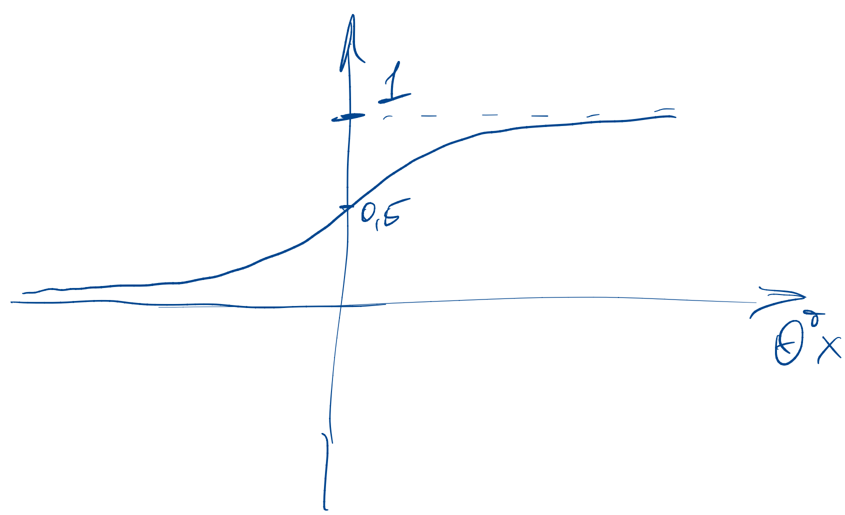
$\mu_i = \Sigma(\theta^T x)$

$p_i = \dfrac{1}{1+e^{-\theta^T x_i}} = \sigma(\theta^T x_i)$

$\mathcal{L}(Y, \theta) =$

$= \prod_{i=1}^{N} p_i =$

$= \prod_{i=1}^{N} \left(p_i^{y_i} \cdot (1-p_i)^{(1-y_i)}\right)$




$l = \ln \mathcal{L} = \sum \ln p_i^{y_i} + \sum \ln (1-p_i)^{(1-y_i)} =$

$= \underline{\sum y_i \ln p_i + \sum (1-y_i) \ln (1-p_i)} =$

$$= \sum_{i=1}^{N} \left( y_i \ln p_i + (1-y_i) \ln (1-p_i) \right) = \ell$$

$$\theta^* = \underset{\Theta}{\arg\max}\ \ell$$

$$\theta^* = \underset{\Theta}{\arg\min} (-\ell)$$

$$L(\mathcal{T}, \theta) = -\sum_{i=1}^{N} \left( y_i \ln p_i + (1-y_i) \ln (1-p_i) \right)$$

$$\nabla_\theta L - ?$$

$$p_i = \frac{1}{1+e^{-\theta^T x_i}}$$

$$\ln \frac{p_i}{1-p_i} = \theta^T x$$

$$1-p_i = \frac{1}{1+e^{\theta^T x}}$$

$$L = \sum \ln (1-p_i) + \sum y_i \left( \ln p_i - \ln (1-p_i) \right) =$$

$$\ln \frac{p_i}{1-p_i} = \theta^T x$$

$$= \sum \ln \frac{1}{1+e^{\theta^T x}} + \sum y_i \left( \theta^T x \right)$$

$$\frac{\partial}{\partial \theta} = y_i x_i$$

$$\frac{\partial}{\partial \theta} \sum \ln (1-p_i) = -\sum p_i x_i$$

$$\nabla_\theta L = \sum \left( y_i - p_i(\theta, x_i) \right) x_i$$