Метод максимального правдоподобия

Maximum Likelihood Estimator

$$P(A,B) = P(A|B)\,P(B) = P(B|A)\,P(A)$$

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

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

$P(\theta)$ – априорное распр-е θ

$P(\theta|\mathcal{T})$ – апостериорное распр-е θ

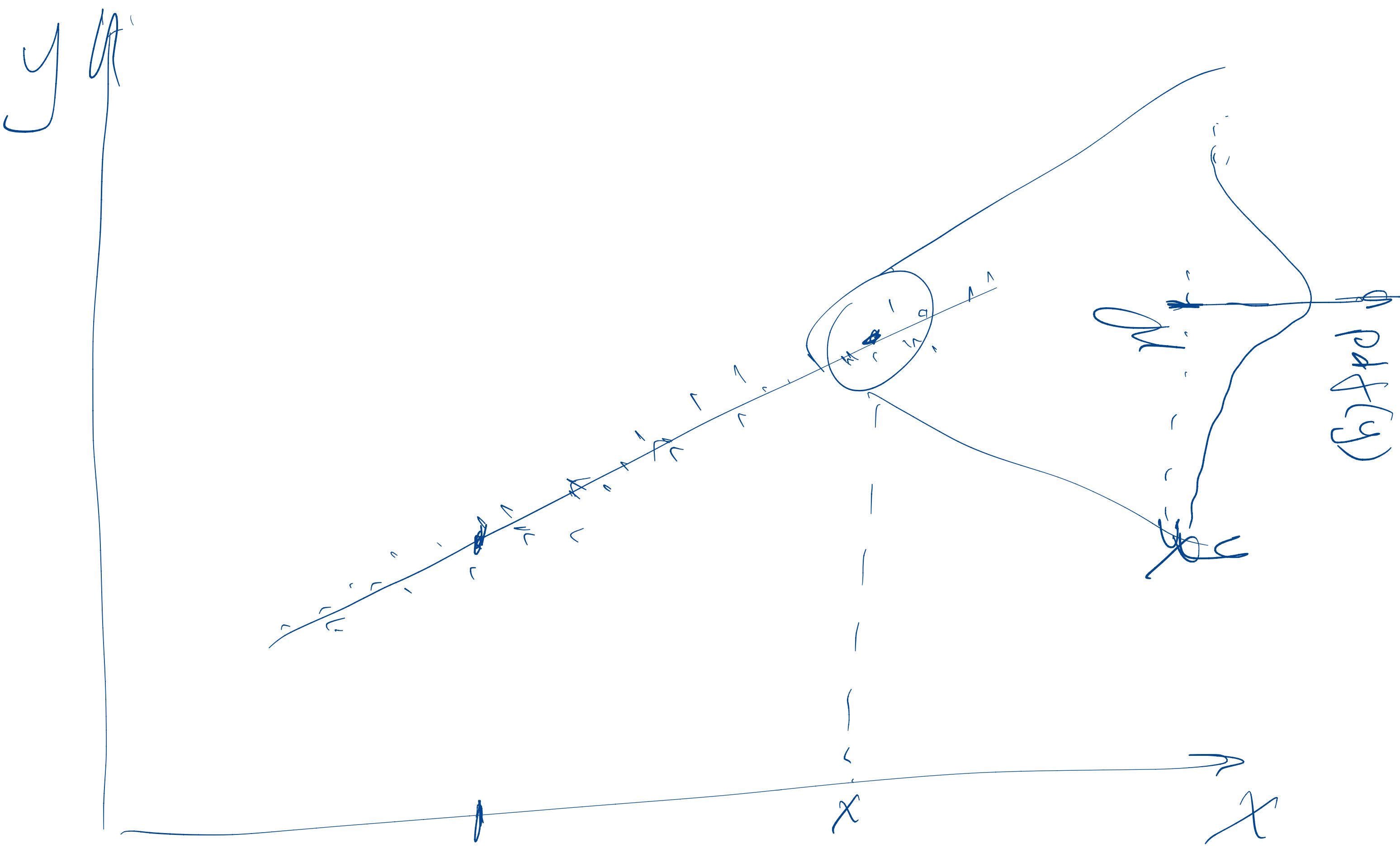
$P(\mathcal{T})$ – evidence, распр-е данных

$P(\mathcal{T}|\theta)$ – правдоподобие выборки $\mathcal{T}$

$$\theta^* = \arg\max_{\Theta} P(\mathcal{T}|\theta)$$

Линейная регрессия

LINE



① L: $\mu = \theta^T X$

② I: $\{X, y\}$: i.i.d.

③ N: $y \sim \mathcal{N}(\theta^T X, \sigma^2)$

④ E: $\sigma_i^2 = \sigma^2$

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

$$L(\mathcal{J}|\theta) = \prod p(x_i, y_i, \theta, \sigma_i) =$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} = ?$$

$$\operatorname{argmax} \ln(L) = \operatorname{argmax} L$$

$$\operatorname*{argmax}_{\Theta} L = \operatorname{argmax} \ln(L) =$$

$$= \operatorname*{argmax}_{\Theta} \left(\sum_{i=1}^{N} \left(\ln \frac{1}{\sqrt{2\pi\sigma_i^2}} + \ln e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} \right) \right)$$

$$= \operatorname*{argmax}_{\Theta} \sum_{i=1}^{N} -\frac{(y_i-\mu_i)^2}{2\sigma^2} = \operatorname*{argmax}_{\Theta} \sum -\frac{(y_i-\mu_i)^2}{2}$$

$$\theta^* = \underset{t+1}{\operatorname{argmax}} \sum -\frac{(y_i - \mu_i)^2}{2} =$$

$$= \underset{t+1}{\operatorname{argmin}} \sum \frac{(y_i - \mu_i)^2}{2} =$$

$$= \underset{t+1}{\operatorname{argmin}} \sum \frac{(y_i - \theta^T x_i)^2}{2}$$