Supervised learning

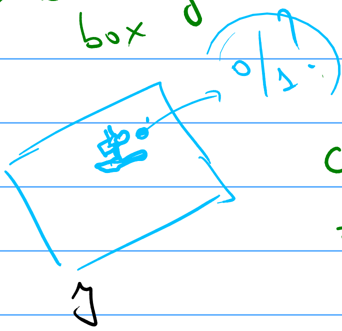
D

$\bar{x} \rightarrow \boxed{} \rightarrow y$

$p(y|\bar{x})$

$$D = \{(\bar{x}_n, y_n)\}_{n=1}^{N}$$

bounding box



$3 \cdot 10^6$

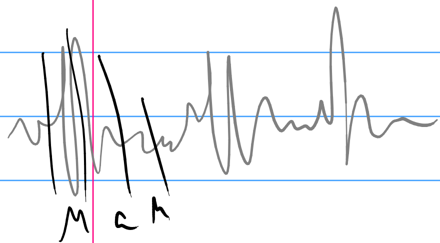
$\bar{x} \in \mathbb{R}$

classification

$\bar{x} \rightarrow c_1, \dots, c_K ?$

regression

$\bar{x} \rightarrow y \in \mathbb{R}$



$\rightarrow$ мама мыла раму

Unsupervised learning

$$D = \{\bar{x}_n\}_{n=1}^{N}$$

clustering

$3 \cdot 10^6$

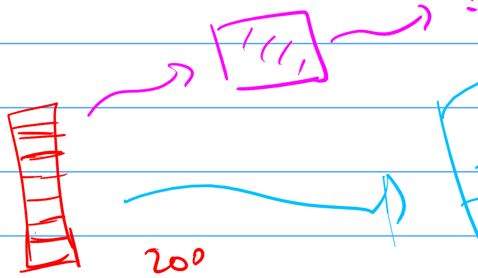
$\mathbb{R}$

$p(\bar{x})$

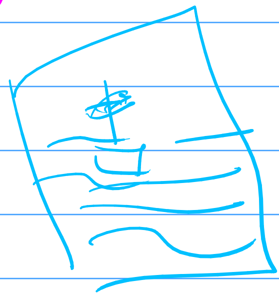
$\bar{x}$

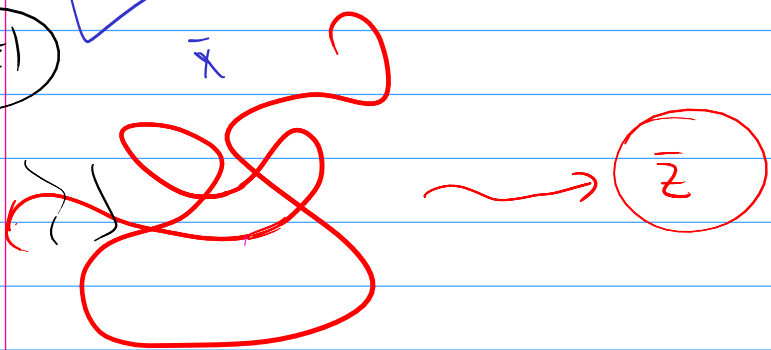
feature extraction

dimensionality reduction



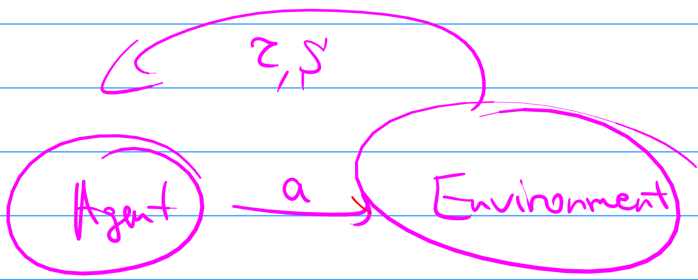
$\bar{z} \in \mathbb{R}^{200}$





$\bar{z}$

Reinforcement learning



$D = \{(\bar{x}, y)\}$

$\bar{x} \;-\; \boxed{\bar{\theta}} \;-\; y$

$\bar{x} \;-\; \boxed{} \;-\; c_1 / c_2 \,?$

prior

$$\boxed{p(\bar{\theta}|D) = \frac{p(D|\bar{\theta})\,p(\bar{\theta})}{p(D)}}$$

posterior

likelihood

$\sum_i p(x = a_i) = 1$

$\int p(x)\,dx = 1$

$\mathcal{N}(x|\bar{\mu}, \Sigma)$

$$p(d|t) = \frac{p(t|d)\,p(d)}{p(t) = p(t|d)\,p(d) + p(t|\bar{d})\,p(\bar{d})}$$

$p(d) = 1/100$

$p(t|d) = 95/100$, $p(d) = 1/100$, $p(t|\bar{d}) = 5/100$, $p(\bar{d}) = 99/100$

$\theta = p(\text{орёл})$

$1-\theta = p(\text{решка})$

$D = hhthtth$

likelihood: $p(D|\theta) = \theta^4 (1-\theta)^3$

$$p(D|\theta) = \theta^n (1-\theta)^m \xrightarrow[\theta]{} \max$$

$$\theta_{ML} = \arg\max_{\theta} p(D|\theta)$$

$$\frac{\partial p(D|\theta)}{\partial \theta} = n\theta^{n-1}(1-\theta)^m - m\theta^n(1-\theta)^{m-1} = 0$$

$$\theta^{n-1}(1-\theta)^{m-1}\big(n(1-\theta) - m\theta\big) = 0$$

$\boxed{\theta = 1, \ \theta = 0}$ $\qquad \boxed{\theta_{ML} = \frac{n}{n+m}}$

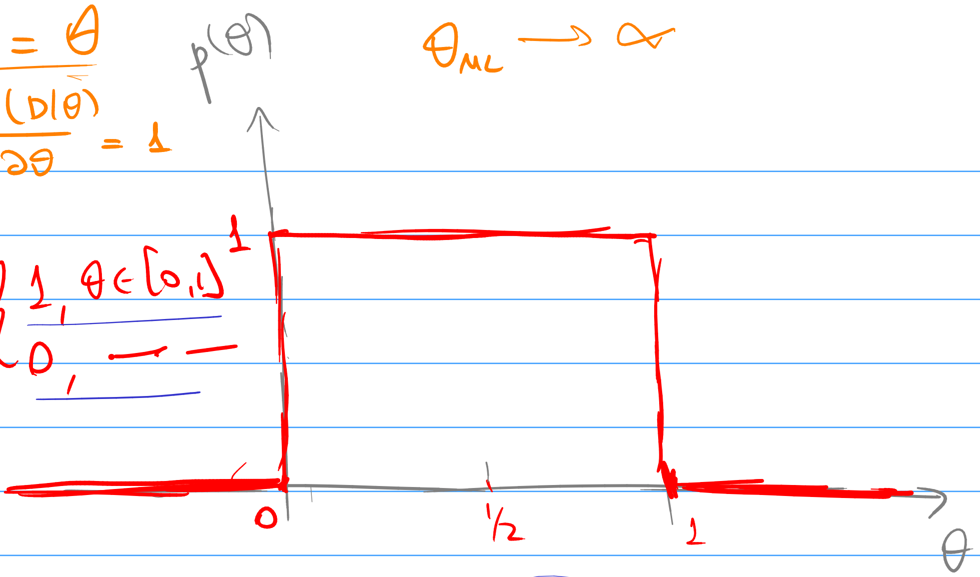
$D = h$

$p(D|\theta) = \theta$

$\frac{\partial p(D|\theta)}{\partial \theta} = 1$

$\theta_{ML} \longrightarrow \alpha$

$p(\theta) = \begin{cases} 1, & \theta \in [0,1] \\ 0, & -- \end{cases}$



posterior: $p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \xrightarrow[\theta]{} \max$

$\theta_{MAP} = \arg\max_{\theta} p(\theta|D)$

$p(\theta|D) = \begin{cases} \frac{1}{p(D)} \cdot \theta^n (1-\theta)^m, & \theta \in [0,1] \\ 0, & \theta \notin [0,1] \end{cases}$

$p(\theta|D) \propto \underline{p(\theta)p(D|\theta)}$

$p(D) = \int p(\theta)p(D|\theta)\,d\theta)$

$\int_0^1 \theta^n (1-\theta)^m d\theta = \frac{\Gamma(n+1)\Gamma(m+1)}{\Gamma(n+m+2)} = \frac{n!\,m!}{(n+m+1)!}$

$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$

$\| $

$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$

$\Gamma(n) = (n-1)!$

$p(\theta|D) = \begin{cases} \frac{(n+m+1)!}{n!\,m!} \theta^n (1-\theta)^m, & \theta \in [0,1] \\ 0, & \theta \notin [0,1] \end{cases}$

$\theta_{MAP} = \frac{n}{n+m}$

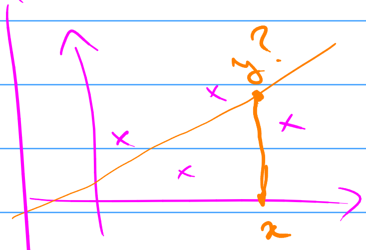
max

θ_{MAP}

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

max

θ_{ML}



$$p(D|\theta) = \prod_{n=1}^{N} p(d_n|\theta)$$

$\{d_1 .. d_N\}$

$$p(\text{heads}|D) = \int p(\text{heads}, \theta|D)\, d\theta =$$

$$= \int p(\text{heads}|\theta, D)\, p(\theta|D)\, d\theta =$$

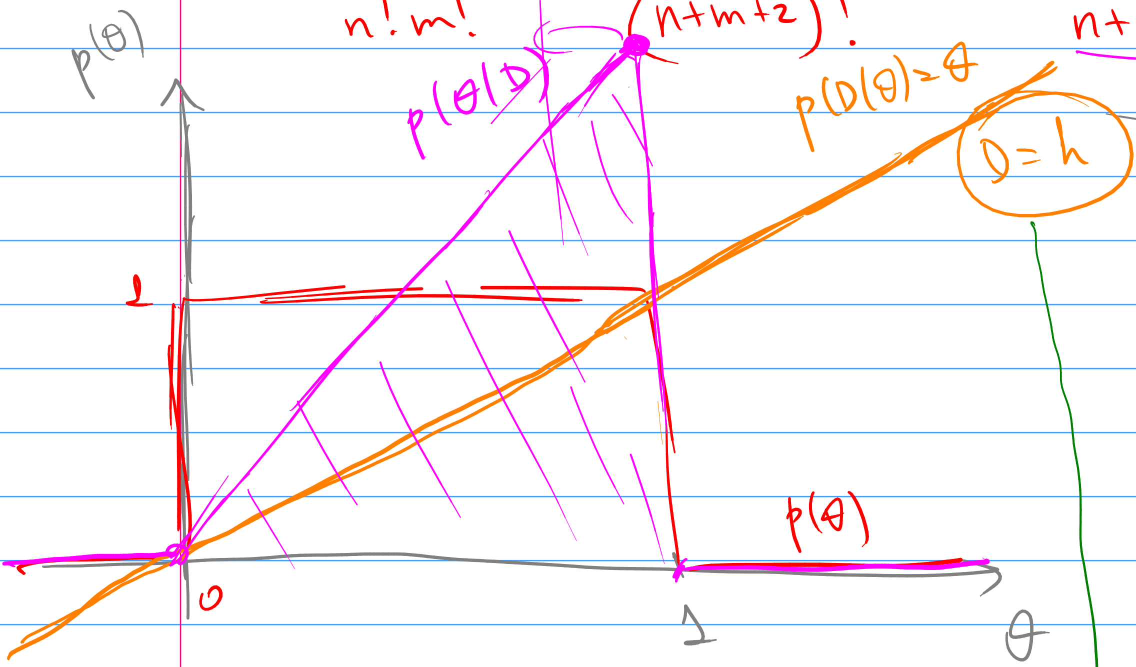
θ posterior

$$= \int p(\text{heads}|\theta)\, p(\theta|D)\, d\theta$$

$$p(\text{heads}|D) = \int \theta \cdot p(\theta|D)\, d\theta = \frac{(n+m+1)!}{n!\,m!} \int_0^1 \theta \cdot \theta^n (1-\theta)^m\, d\theta =$$

$$= \frac{(n+m+1)!}{n!\,m!} \cdot \frac{(n+1)!\,m!}{(n+m+2)!} = \frac{n+1}{n+m+2}$$

Laplace's rule of succession



$$p(x|D) = \int p(x, \theta|D)\, d\theta = \int p(x|\theta)\, p(\theta|D)\, d\theta = \mathbb{E}_{\theta \sim p(\theta|D)}\left[p(x|\theta)\right]$$

$$p(y|D, \bar{x}) = \int p(y|\bar{x}, \theta)\, p(\theta|D)\, d\theta = \mathbb{E}_{p(\theta|D)}\left[p(y|\bar{x}, \theta)\right]$$