

MLE vs. MAP

$$\begin{aligned}\theta_{MLE} &= \operatorname{argmax} P(x|\theta) \\ &= \operatorname{argmax} \prod_i P(x_i|\theta) \\ &= \operatorname{argmax} \sum_i \log p(x_i|\theta)\end{aligned}$$

$$P(\theta|x) = \frac{P(x|\theta) P(\theta)}{P(x)} \propto P(x|\theta) P(\theta)$$

$$\begin{aligned}\theta_{MAP} &= \operatorname{argmax} \{P(x|\theta) P(\theta)\} \\ &= \operatorname{argmax} \{\log(P(x|\theta) P(\theta))\} \\ &= \operatorname{argmax} \{\log P(x|\theta) + \log P(\theta)\} \\ &= \operatorname{argmax} \{\log \prod_i P(x_i|\theta) + \log P(\theta)\} \\ &= \operatorname{argmax} \{\sum_i \log p(x_i|\theta) + \log p(\theta)\}\end{aligned}$$

prior

$$\begin{aligned}P(Y|x) &= \frac{P(x|Y) P(Y)}{P(x)} \\ &\propto P(x|Y) P(Y)\end{aligned}$$

posterior \swarrow likelihood \swarrow prior \swarrow evidence