MLE vs. MAP

$$\theta_{MLE} = \operatorname{argmax} P(X|\theta)$$

$$= \operatorname{argmax} \prod_i P(x_i|\theta)$$

$$= \boxed{\operatorname{argmax} \sum_i \log p(x_i|\theta)}$$

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

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \propto P(X|Y)P(Y)$$

evidence

posterior

likelihood

prior

$$\theta_{MAP} = \operatorname{argmax}\{P(X|\theta)P(\theta)\}$$

$$= \operatorname{argmax}\{\log(P(X|\theta)P(\theta))\}$$

$$= \operatorname{argmax}\{\log P(X|\theta) + \log P(\theta)\}$$

$$= \operatorname{argmax}\{\log \prod_i P(x_i|\theta) + \log p(\theta)\}$$

$$= \boxed{\operatorname{argmax}\{\sum_i \log p(x_i|\theta) + \log p(\theta)\}}$$

prior