

1. 机器学习开篇

P1 频率派 vs. 贝叶斯派

X : data.

θ : parameter

$X \sim P(X|\theta)$.

频率派: θ : 标定的常量 X : random variable.

MLE: $\theta_{MLE} = \arg\max_{\theta} \log P(X|\theta)$

统计机器学习.
 ① 模型
 ② loss
 ③ 优化问题

贝叶斯派: θ : random variable. , $\theta \sim p(\theta)$ 先验.

$$P(\theta|X) = \frac{\overset{\text{likelihood}}{P(X|\theta)} \cdot \overset{\text{prior}}{P(\theta)}}{\underset{\text{posterior}}{P(X)}} \propto P(X|\theta) \cdot P(\theta).$$

$$\int_{\theta} P(X|\theta) \cdot P(\theta) d\theta.$$

贝叶斯 \rightarrow 概率图模型.
 求积分问题 (MCMC)

MAP: $\theta_{MAP} = \arg\max_{\theta} P(\theta|X) = \arg\max_{\theta} P(X|\theta) \cdot P(\theta).$



贝叶斯估计: $P(\theta|X) = \frac{P(X|\theta) \cdot P(\theta)}{\int_{\theta} P(X|\theta) \cdot P(\theta) d\theta}$

贝叶斯预测: $X, \tilde{x}, P(\tilde{x}|X) = \int_{\theta} P(\tilde{x}, \theta|X) d\theta.$

$X \rightarrow \theta \rightarrow \tilde{x} \quad = \int_{\theta} P(\tilde{x}|\theta) P(\theta|X) d\theta.$