

极大似然估计 vs 最大后验概率估计

MLP

MAP

MLP: $X_i, i=1, \dots, n$ 是独立采样于某分布的数据, 估计参数 θ , 使得在这些参数下观测事件发生的概率最大。

$$L(x_1, \dots, x_n | \theta) = \prod_{i=1}^n P(x_i | \theta) \Rightarrow \max_{\theta} L(x_1, \dots, x_n | \theta)$$

example: 抛一枚硬币三次, 得 0, 1, 1, 1 代表正面, 0 代表反面

请估计正面朝上的概率 θ .

$$\max_{\theta} \theta^2 (1-\theta) \Rightarrow \frac{\partial \theta^2 (1-\theta)}{\partial \theta} = 0 \Rightarrow \hat{\theta} = \frac{2}{3}$$

MAP: 在 MLP 中, θ 被视为常数, 但也可以把 θ 认为是一个变量, 具有一个先验分布 $p(\theta)$, 从而

$$\max_{\theta} \prod_{i=1}^n P(x_i | \theta) p(\theta), \quad p(\theta) \text{ 是已知的函数.}$$