

# chapter 2: 贝叶斯网络

## 符号介绍

同 chapter 1

## 概念

- MLE
- MAP
- conjugate distribution

## 知识点

Q1: 先验概率的基本意义  
A1: 相当于直接给模型灌输"规则(expert knowledge)" ps: L1正则化相当于给 $\theta$ 添加了Laplace先验分布, L2正则化相当于给 $\theta$ 增加了Gussian先验分布

Q2: 简述共轭分布  
A2: 对于给定的 $P(X|\theta)$ , 以其为似然函数, 必然存在某一分布的先验概率 $P(\theta)$ 与后验概率 $P(\theta|X)$ 具有同一形式, 此时 $P(X|\theta)$ 与 $P(\theta|X)$ 这两者的分布互为共轭分布

Q3: 简述NB训练过程  
A3:

$$J(\theta) = \operatorname{argmax} P(Y = y|X = x) = \operatorname{argmax} P(X = x|Y = y)P(Y = y)$$

朴素贝叶斯假设数据 $x$ 的所有属性相互独立, 则原式可化为:

$$= \operatorname{argmax} \prod_{i=1}^{i=m} P(X_i = x_i|Y = y)P(Y = y)$$

而 $P(x_i|y)$ 与 $P(y)$ 两者都可以通过统计训练数据中的频率得到

- 
- 独立性假设过于严格, 但是在实践过程中即便属性之间不遵循相互独立原则, 分类效果依旧可圈可点.
  - 对于统计概率中出现的0概率情况的解决办法, 采用laplacian correction
  - 对于连续值的属性可以采用Gaussian distribution

Q4: 简述生成式模型与判别式模型的区别  
A4: 判别式模型是直接对 $p(y|x)$ 进行建模, 而生成式子模型通过对 $P(x, y)$ 联合概率或者采用bayes公式对 $p(x|y)p(y)$ 进行建模. 常见的判别式模型主要有LR, SVM, NN; 常见的生成式模型主要有: Bayes classifier, HMM, GMM

## 公式推导以及证明