DATE

概率图1— 背景介绍

1. 高维随机变量计算量大 → 简化 $\xrightarrow{\text{相互独立}}$ $p(x_1,\cdots,x_p)=\prod_{j=1}^{p} p(x_j)$

Naive Bayes: $P(X|y)=\prod_{j=1}^{p} p(x_j|y)$

Markov Property → $x_j \perp x_{i+1} | x_i,\ j<i$

HMM: 齐次Markov假设

条件独立性 → $x_A \perp x_B | x_C$, x_A, x_B, x_C是集合且不相交 (是马尔可夫性质的推广)

2. 概率图
- representation
  - 有向图 Bayesian Network
  - 无向图 Markov Network
- inference 推断
  - 精确推断
  - 近似推断
    - 确定性近似 (变分推断)
    - 随机近似 (MCMC)
- learning 学习
  - 结构学习
  - 参数学习
    - 完备数据
      - 有向图
      - 无向图
    - 隐变量: EM
- decision 决策

概率图2 — 贝叶斯网络

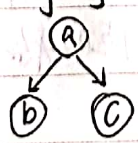
1. chain rule: $P(x_1 \cdots x_p)=p(x_1)\prod_{i=2}^{p}(x_i|x_{1:i-1})$

2. 因子分解 $P(x_{1:p})=\prod_{i=1}^{p} p(x_i|x_{pa(i)})$

$x_{pa(i)}$是x_i的父节点集合

3. 拓扑排序: $i<j \longrightarrow [\cdots, x_i, \cdots, x_j \cdots]$

4. 假定已构造了一个有向图模型:



$p(a,b,c)=\underline{p(a)\,p(b|a)\,p(c|a)}$ 因子分解

$p(a,b,c)=\underline{p(a)\,p(b|a)\,p(c|a,b)}$ 链式法则

所以 $p(c|a)=p(c|a,b) \Rightarrow$ ~~b,c独立~~ $b \perp c | a$

证明: $p(c|a)p(b|a)=p(c|a,b)p(b|a)$

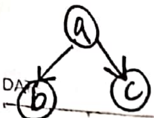
$=p(c,b|a)$

说明图与集合都能体现独立性.

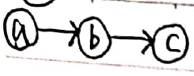
$c \perp b | a \Leftrightarrow$ a若被观测, 则路径阻塞

(无向)路径: $b \leftarrow a \rightarrow c$ (tail to tail) tail: a是头b是尾

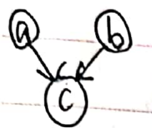
有向图三种情况



tail to tail　$b \perp c \mid a$



head to tail　$a \perp c \mid b$ ⟺ 若b被观测则路径被阻塞



head to head　默认情况下 $a \perp b$，路径堵塞

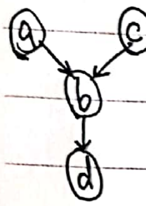
☆ 若c被观测，a和b不独立，路径是通的　　类比：两个子结婚生娃C

证明：$P(a,b,c) = P(a)P(b)P(c|a,b)$

$= P(a)P(b|a)P(c|a,b)$

所以 $P(b) = P(b|a) \Rightarrow a \perp b$

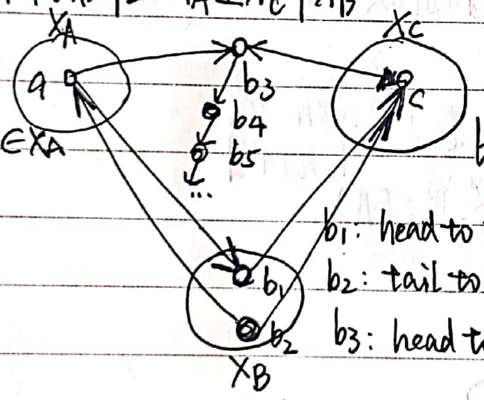
概率图3——D划分　D-Seperation



head to head的拓展

若d被观测，a-b-c也是连通的

条件独立性　$X_A \perp X_C \mid X_B$



$\forall a \in X_A$　　$\forall c \in X_C$

D划分（全局马尔可夫性）

b_1, b_2 一定在 X_B 中

$b_3, b_4, b_5, \cdots$ 一定不在 X_B 中

b_1: head to tail

b_2: tail to tail

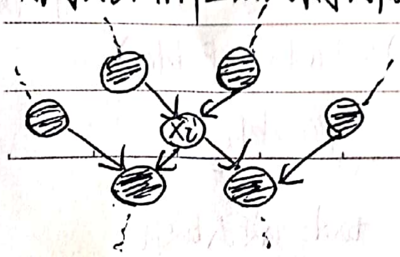
b_3: head to head

$$P(x_i \mid x_{-i}) = \frac{P(x_i, x_{-i})}{P(x_{-i})} = \frac{P(x)}{\underbrace{\int_{x_i} P(x)\,dx_i}_{\text{边缘分布转联合分布}}} = \frac{\prod_{j=1}^{p} P(x_j \mid x_{pa(j)})}{\int_{x_i} \prod_{j=1}^{p} P(x_j \mid x_{pa(j)})\,dx_i}$$

x_{-i} 表示集合 $\{x_1 \ldots i-1, i+1, \ldots, p\}$

$= f(\triangle)$，用△表示与 x_i 相关的点的集合（未给出证明）

i.e. x_i 与其它所有节点的关系可以简化为 x_i 与“和 x_i ~~连通~~ 相关的点集”的关系。

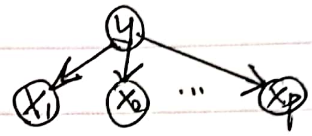


Markov Blanket　马尔可夫毯

概率图4 — representation example

贝叶斯网络

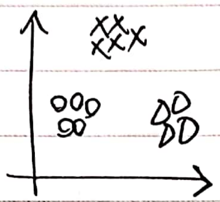
Naive Bayes: $P(X|y)=\prod_{i=1}^{p} p(x_i|y=1)$



GMM: $Z \rightarrow X$

Z is discrete

i.e. $z=1,2,\dots,K$



$X|Z \sim N(\mu,\Sigma)$

带时间: Markov Chain, 随机过程中的一种

Gaussian Process 无限维高斯分布

连续: Gaussian Network / Gaussian Bayesian Network

动态系统

e.g. HMM

离散隐状态序列是马氏链

e.g. LDS

卡尔曼滤波

线性,连续(高斯)

e.g. Particle Filter

非高斯,非线性