第11章 概率图模型

11.1 模型表示

带阴影的节点表示可观测的变量.

不带阴影的节点表示隐含变量.

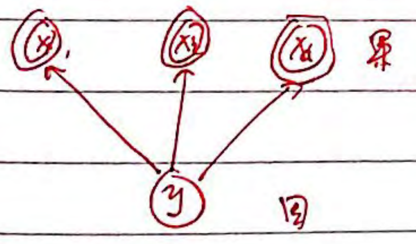
11.1.1 有向图模型（贝叶斯网络）

11.1.2 常见有向图模型

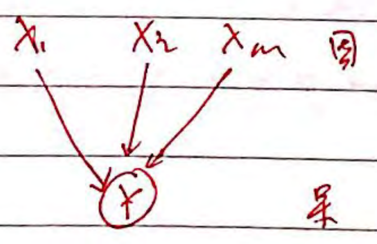
Sigmoid 信念网络

11.1.3 朴素贝叶斯分类器

$$p(y|x,\theta) \propto p(y|\theta_c)\prod_{i=1}^{d} p(x_i|y,\theta_{i|y})$$

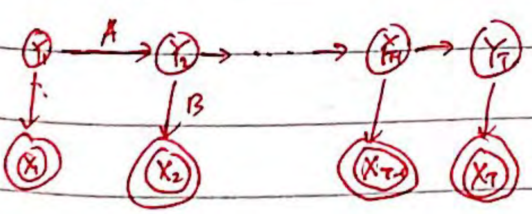


朴素贝叶斯



Logistic 回归.

11.2.3 HMM



$$p(x,y,\theta) = \prod_{t=1}^{T} \underbrace{p(y_t|y_{t-1},\theta)}_{A}\ \underbrace{p(x_t|y_t,\theta_t)}_{B}$$

11.1.3 无向图模型（马尔科夫随机场）

11.1.4 无向图模型的概率分解

11.1.5 常见的无向图模型.

对数线性模型.（最大熵模型）

条件随机场 CRF

11.1.5.1 对数线性模型.

$f_c(x_c)$ 特征向量.

θ_c 权重向量.

势能函数

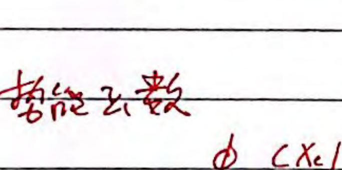
$$\phi_c(x_c|\theta_c)=\exp(\theta_c^T f_c(x_c))$$

对数线性模型

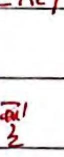
$$\log p(x|\theta)=\sum \theta^T f_c(x_c) - \log Z(\theta)$$

$$p(y|x,\theta)=\frac{1}{Z(x,\theta)}\exp(\theta^T f(x,y))$$

11.1.5.2 CRF



$$p(y|x)=\frac{1}{Z(x,\theta)}\exp(\sum\theta_c^T f_c(x,y_c))$$



11.2 推断

11.3 近似推断

11.3.1 采样法（蒙特卡罗方法）

计算函数 $f(x)$ 的期望，已知 $p(x)$

$$E(f(x)) = \int_x f(x)p(x)\,dx$$

随机采样

11.3.2 拒绝采样

引入一个容易采样的分布 $q(x)$ —— 提议分布

构建一个提议分布 $q(x)$ 和常数 k，使得 $kq(x)$ 可以覆盖函数 $\hat{p}(x)$

即 $kq(x) \ge \hat{p}(x)$

对于每次抽取的样本 $\hat{x}$，计算接受概率 $\alpha(\hat{x})$，并以 $\alpha(\hat{x})$ 来接受样本 $\hat{x}$.

$$\alpha(\hat{x}) = \frac{\hat{p}(\hat{x})}{kq(\hat{x})}$$

No.

Date.

11.3.3 重要性采样

$$E_p[f(x)] = \int_x f(x)p(x)dx$$

$$= \int_x f(x) \frac{p(x)}{q(x)} q(x) dx$$

$$= \int_x f(x)\, w(x)\, q(x)\, dx$$

$$= E_q [f(x) w(x)].$$

$w(x)$ 重要性权重.

将分布 $p(x)$ 在 $f(x)$ 的期望变为分布 $q(x)$ 在 $f(x)w(x)$ 的期望.

11.3.4 MCMC

高维空间采样

- MH算法

- Metropolis算法

- 吉布斯采样

No.

Date. . .

11.6 学习

11.6.1 不含隐变量的参数估计

有向图 极大似然估计

无向图 采样/坐标上升法

11.6.2 含隐变量的参数估计

11.6.2.1 EM算法

一个样本的边际似然函数

$$p(x|\theta)=\sum_{z}p(x,z|\theta)$$

带隐变量的贝叶斯网络