- 概率图模型
  - 是什么，为什么需要概率图模型。
- 基本问题
  - 模型表示
    - 有向图，无向图，道德化，盘式记法。
    - 团，极大团，全局、局部、部分马尔科夫性，分离…
  - 学习和推断。
    - 最大似然，EM算法。
    - 精确推断：变量消去、信念传播
    - 近似推断：MCMC采样（随机近似）、变分推断（确定性近似）
- 具体模型
  - 隐马尔科夫模型
  - 条件随机场
- 话题模型

（本章内容比较多，且与第七章关联较大）

> 复习必须看第七章

一、引子

- 生成式：计算联合分布 $P(Y,R,O)$
- 判别式：计算条件分布 $P(Y,R|O)$

□ 符号约定

- Y为关心的变量的集合，O为可观测变量集合，R为其他变量集合

推断inference

> 利用已知变量推测未知变量的条件分布

- 推断的目的是通过生成式或判别式模型，得到条件概率分布P（Y|O）

概率图模型

> 一类用图来表达变量相关关系的概率模型

直接利用概率求和规则消去变量R的时间和空间复杂度为**指数级别**$O(2^{|Y|+|R|})$，需要一种能够简洁紧凑表达变量间关系的工具

- 图模型提供了一种**描述框架**，
  - 结点：随机变量（集合）
  - 边：变量之间的依赖关系

- 分类：
  - 有向图：贝叶斯网
    - 使用有向无环图表示变量之间的依赖关系
  - 无向图：马尔可夫网
    - 使用无向图表示变量间的相关关系

有向图能分析出独立性关系

二、隐马尔可夫模型

> Hidden Markov Model，HMM结构最简单的贝叶斯网

状态变量是隐藏的，表示所有时刻系统状态，下一时刻状态只与上一时刻真实状态有关

观测变量表示所有时刻的观测值，只与当前时刻的隐变量有关

所有状态联合分布：

$$P(x_1, y_1, \ldots, x_n, y_n) = P(y_1)P(x_1 \mid y_1)\prod_{i=2}^{n} P(y_i \mid y_{i-1})P(x_i \mid y_i)\ .$$

生成序列：

通过指定状态空间 $\mathcal{Y}$、观测空间 $\mathcal{X}$ 和上述三组参数，就能确定一个隐马尔可夫模型，通常用其参数 $\lambda = [\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}]$ 来指代. 给定隐马尔可夫模型 λ，它按如下过程产生观测序列 $\{x_1, x_2, \ldots, x_n\}$：

(1) 设置 $t = 1$，并根据初始状态概率 $\boldsymbol{\pi}$ 选择初始状态 y_1；

(2) 根据状态 y_t 和输出观测概率 $\mathbf{B}$ 选择观测变量取值 x_t；

(3) 根据状态 y_t 和状态转移矩阵 $\mathbf{A}$ 转移模型状态，即确定 y_{t+1}；

(4) 若 $t < n$，设置 $t = t + 1$，并转到第 (2) 步，否则停止.

三、马尔可夫随机场

> Markov Random Field，MRF
> 典型的马尔可夫网，无向图

势函数（因子），用于定义概率分布函数

团：全连接子图

极大图：无法再添加点以保持全连接性（无法被其他团包含，可以分解成全部都是极大团）

多个变量之间的联合概率分布能基于团分解为多个因子的乘积，每个因子仅与一个团相关

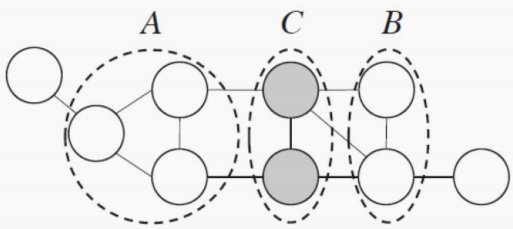
找极大团与概率无关，难

- 基于**极大团**的势函数：
  - 通过极大团构造势函数。若团Q不是一个极大团，则必然被一个极大团Q^*包含，这意味着变量$\mathbf{x}_Q$的关系不仅体现在势函数ψ_Q中，还体现在ψ_{Q^*}中
  - 联合概率分布可以使用极大团定义
  - 假设所有极大团构成的集合为$\mathcal{C}^*$

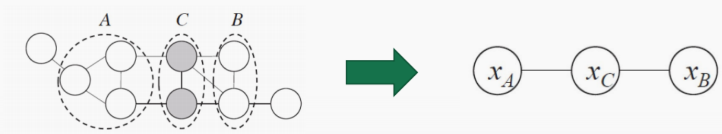
$$P(\mathbf{x}) = \frac{1}{Z^*} \prod_{Q \in \mathcal{C}^*} \psi_Q(\mathbf{x}_Q)$$

  - 其中，Z^*是规范化因子 $Z^* = \sum_{\mathbf{x}} \prod_{Q \in \mathcal{C}^*} \psi_Q(\mathbf{x}_Q)$

- 借助“分离”的概念，若从结点集A中的结点到B中的结点都必须经过结点集C中的结点，则称结点集A，B被结点集C分离，称C为分离集（separating set）



- **全局**马尔可夫性（global Markov property）：在给定**分离集**的条件下，两个变量子集条件独立
  - 若令A,B,C对应的变量集分别为$\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C$，则$\mathbf{x}_A$和$\mathbf{x}_B$在$\mathbf{x}_C$给定的条件下独立，记为$\mathbf{x}_A \perp \mathbf{x}_B \mid \mathbf{x}_C$
- 图模型简化：



- 得到图模型的联合概率为：

$$P(x_A, x_B, x_C) = \frac{1}{Z}\psi_{AC}(x_A, x_C)\psi_{BC}(x_B, x_C)$$

具象化思维不能帮助理解抽象化本质

GMM

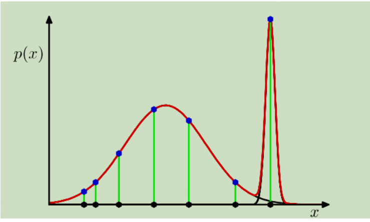
> Kmeans进阶版

两部的生成式模型：第一步确定z，第二部采样x，得到P（x，z）联合分布

likelihood函数

高斯的似然是二次函数

singular fit畸形拟合



因为样本点自己就是一个高斯分布时最大，不能让样本自己独立

EM：期望+最大化

平均场

z不能一味选最简单的，还要考虑事实需求

z拆成M份

zj服从的最优分布与假设无关

> 优秀分布，需要你调的参数很少

[[折中的问题折中解决：q（z）怎么来的?