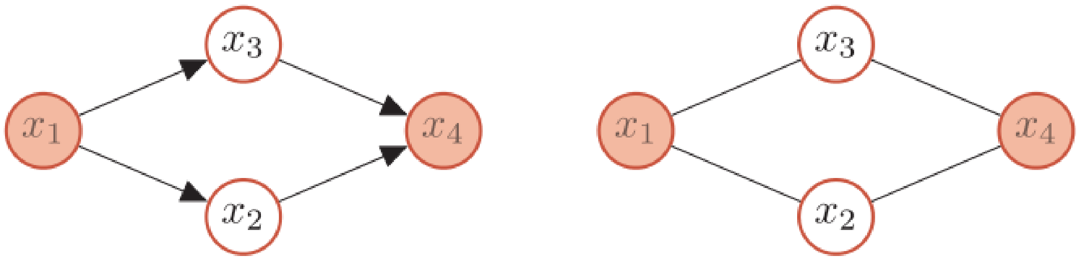
《神经网络与深度学习》

概率图模型

https://nndl.github.io/

概率图模型

- 概率图模型是指一种用图结构来描述多元随机变量之间条件独立关系的概率模型。

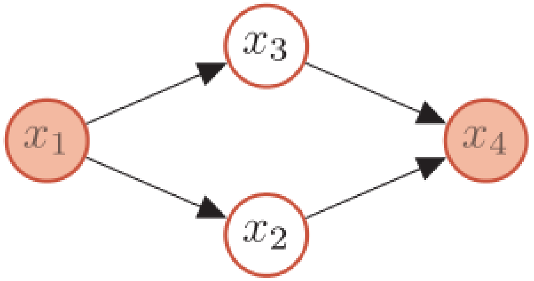


(a) 有向图：贝叶斯网络　　(b) 无向图：马尔可夫随机场

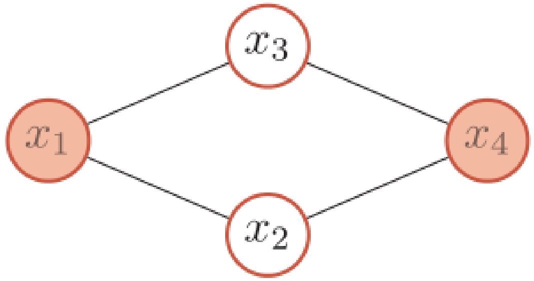
- 每个节点都对应一个随机变量，可以是观察变量，隐变量或是未知参数等；
- 每个连接表示两个随机变量之间具有依赖关系。

概率图模型

- 模型表示（图结构）
  - 有向图
  - 无向图



(a) 有向图：贝叶斯网络

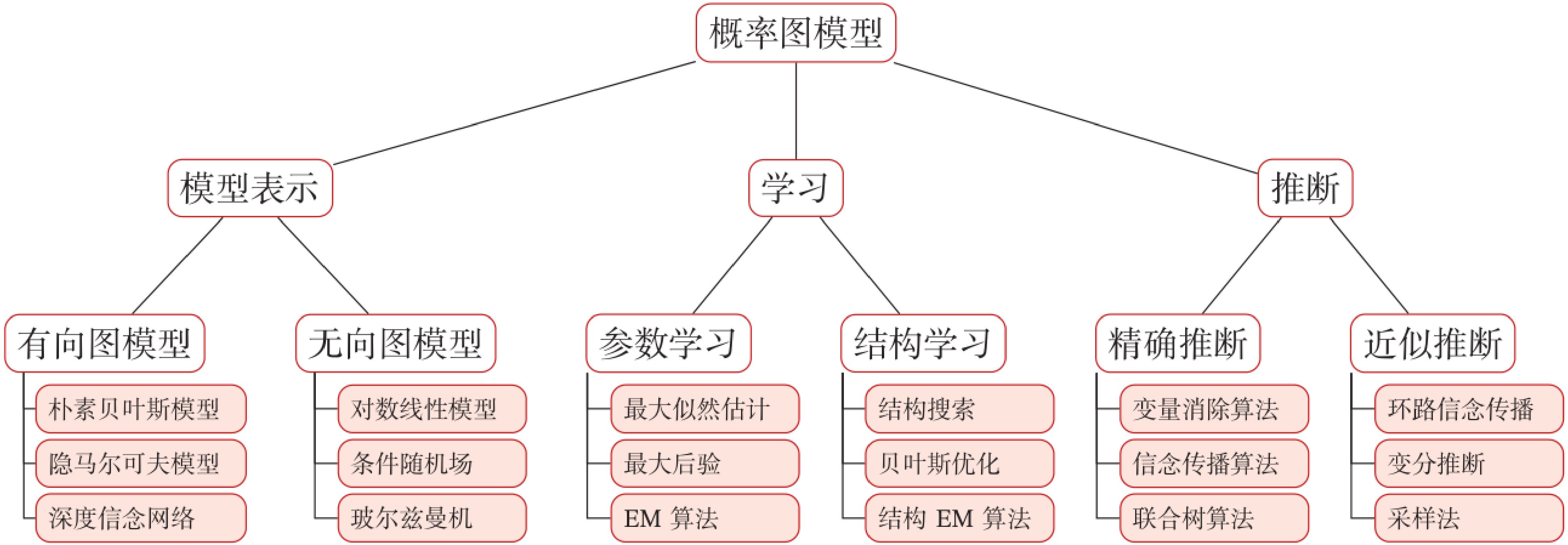


(b) 无向图：马尔可夫随机场

- 推断（ Inference ）
  - 给定部分变量，推断另一部分变量的后验概率。

- 学习（Learning）
  - 参数学习：给定一组训练样本，求解模型参数

概率图模型



概率图模型
模型表示
学习
推断
有向图模型
朴素贝叶斯模型
隐马尔可夫模型
深度信念网络
无向图模型
对数线性模型
条件随机场
玻尔兹曼机
参数学习
最大似然估计
最大后验
EM 算法
结构学习
结构搜索
贝叶斯优化
结构 EM 算法
精确推断
变量消除算法
信念传播算法
联合树算法
近似推断
环路信念传播
变分推断
采样法

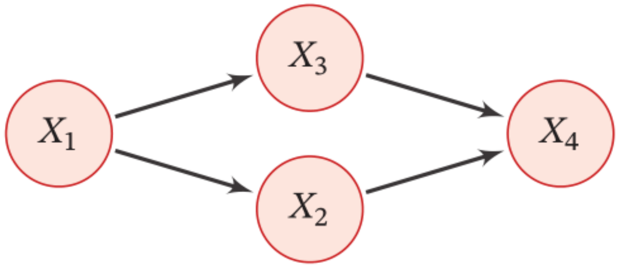
贝叶斯网络

- 有向图模型（Directed Graphical model），也称为贝叶斯网络（Bayesian Network），或信念网络（Belief Network，BN）。

定义 11.1 – 贝叶斯网络： 对于一个 K 维随机向量 $\boldsymbol{X}$ 和一个有 K 个节点的有向非循环图 G，G 中的每个节点都对应一个随机变量，每个连接 e_{ij} 表示两个随机变量 X_i 和 X_j 之间具有非独立的因果关系. 令 $\boldsymbol{X}_{\pi_k}$ 表示变量 X_k 的所有父节点变量集合，$P(X_k|\boldsymbol{X}_{\pi_k})$ 表示每个随机变量的局部条件概率分布（Local Conditional Probability Distribution）. 如果 $\boldsymbol{X}$ 的联合概率分布可以分解为每个随机变量 X_k 的局部条件概率的连乘形式，即

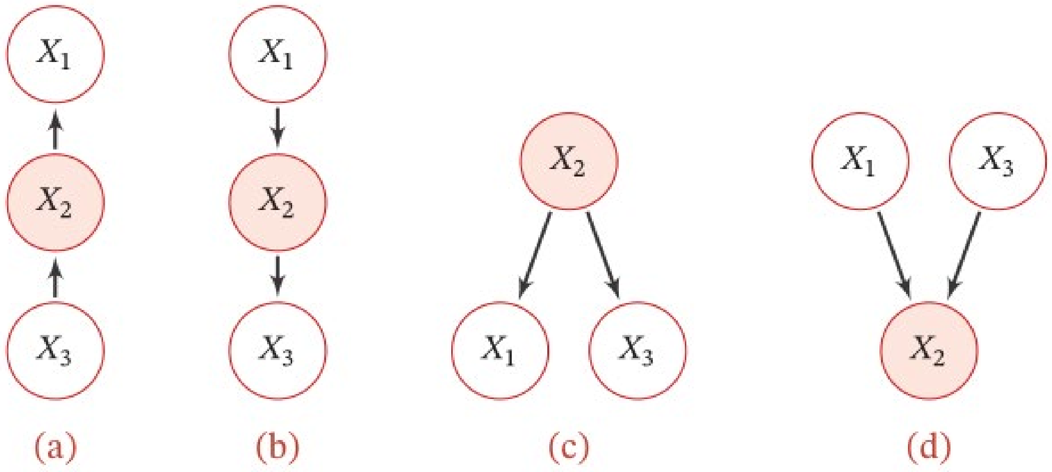
$$p(\boldsymbol{x}) = \prod_{k=1}^{K} p(x_k|\boldsymbol{x}_{\pi_k}), \tag{11.8}$$

那么 $(G, \boldsymbol{X})$ 构成了一个贝叶斯网络.



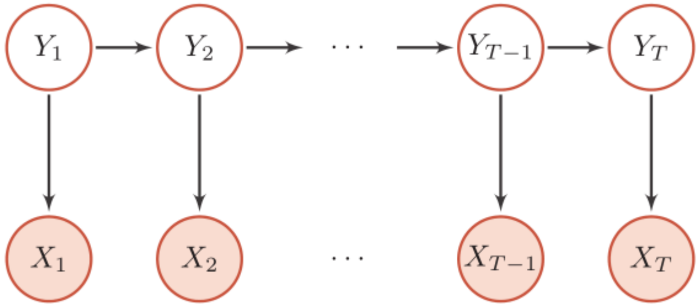
条件独立性

- 在贝叶斯网络中，如果两个节点是直接连接的，它们肯定是非条件独立的，是直接因果关系。
- 父节点是“因”，子节点是“果”。
- 如果两个节点不是直接连接的，但是它们之间有一条经过其他节点的路径连接互连接，它们之间的条件独立性就比较复杂。



隐马尔可夫模型（Hidden Markov Model，HMM）

- 表示一种含有隐变量的马尔可夫过程



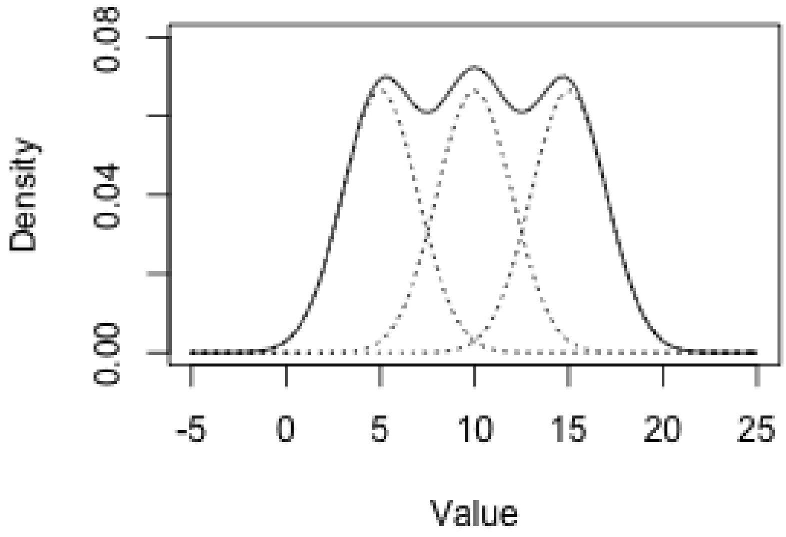
- 隐马尔可夫模型的联合概率可以分解为

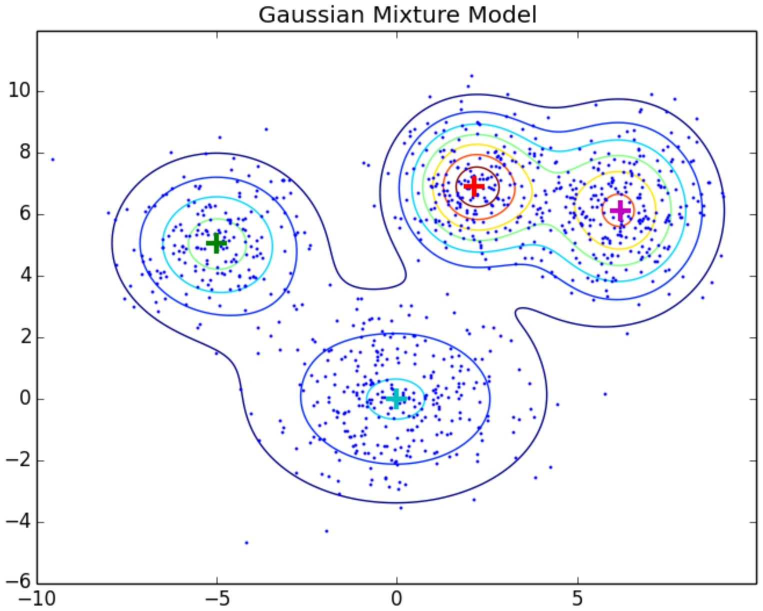
$$p(\boldsymbol{x}, \boldsymbol{y}; \theta) = \prod_{t=1}^{T} p(y_t | y_{t-1}, \theta_s) p(x_t | y_t, \theta_t).$$

转移概率　输出概率

高斯混合模型

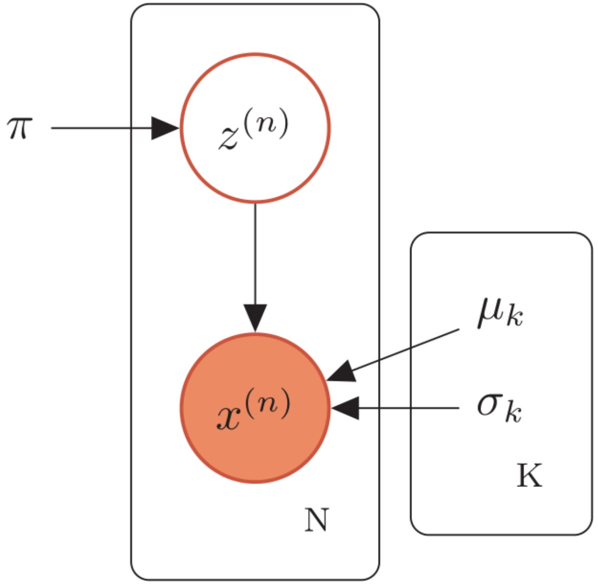
- 高斯混合模型（Gaussian Mixture Model，GMM）是由多个高斯分布组成的模型，其密度函数为多个高斯密度函数的加权组合。





高斯混合模型

图模型表示



$$\mathcal{N}(x|\mu_k, \sigma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_k^2}\right)$$

马尔可夫随机场

- 马尔可夫随机场，也称无向图模型，是一类用无向图来表示一组具有马尔可夫性质的随机变量X的联合概率分布模型。

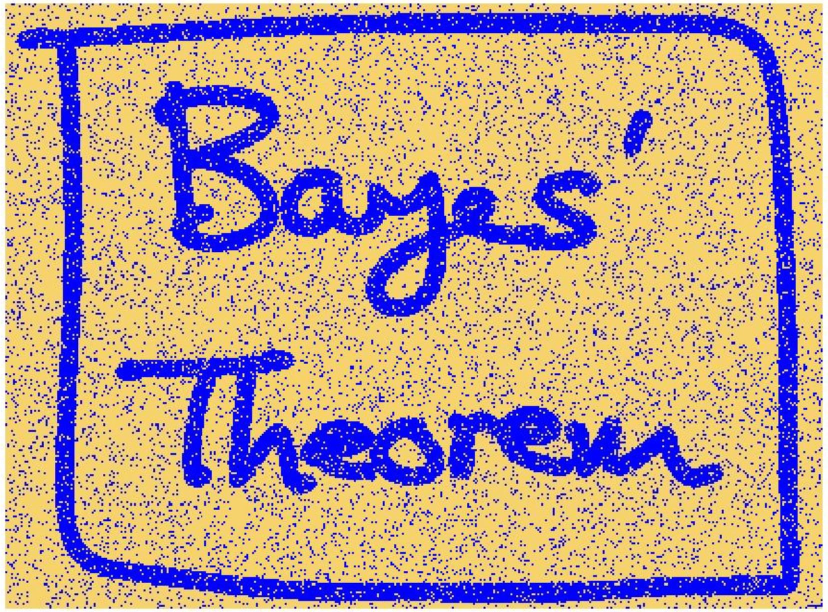
定义 11.2 – 马尔可夫随机场：对于一个随机向量 $\boldsymbol{X} = [X_1, \cdots, X_K]^{\mathrm{T}}$ 和一个有 K 个节点的无向图 $G(\mathcal{V}, \mathcal{E})$（可以存在循环），图 G 中的节点 k 表示随机变量 X_k，$1 \leq k \leq K$. 如果 $(G, \boldsymbol{X})$ 满足局部马尔可夫性质，即一个变量 X_k 在给定它的邻居的情况下独立于所有其他变量，

$$p(x_k|\boldsymbol{x}_{\setminus k}) = p(x_k|\boldsymbol{x}_{\mathcal{N}(k)}), \tag{11.15}$$

其中 $\mathcal{N}(k)$ 为变量 X_k 的邻居集合，$\setminus k$ 为除 X_k 外其他变量的集合，那么 $(G, \boldsymbol{X})$ 就构成了一个马尔可夫随机场.

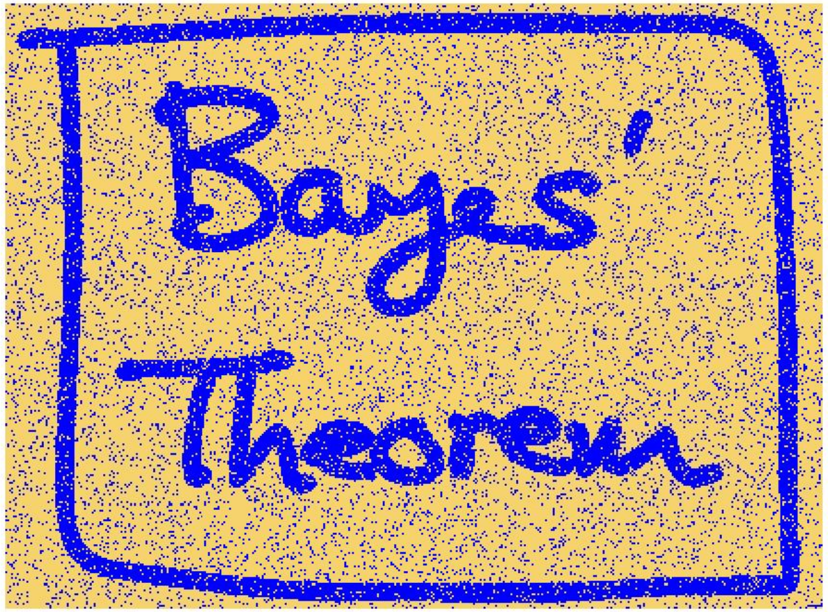
Illustration: Image De-Noising

Original Image

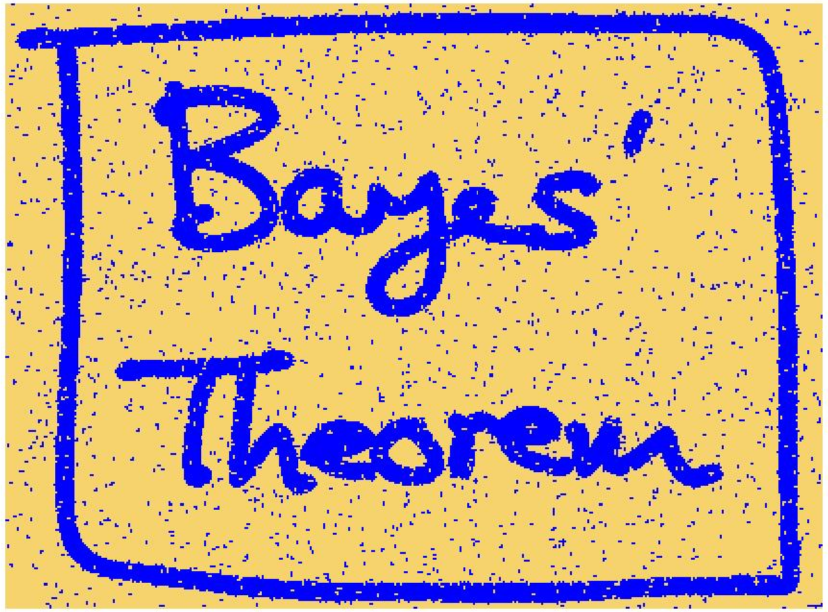


Noisy Image

Illustration: Image De-Noising

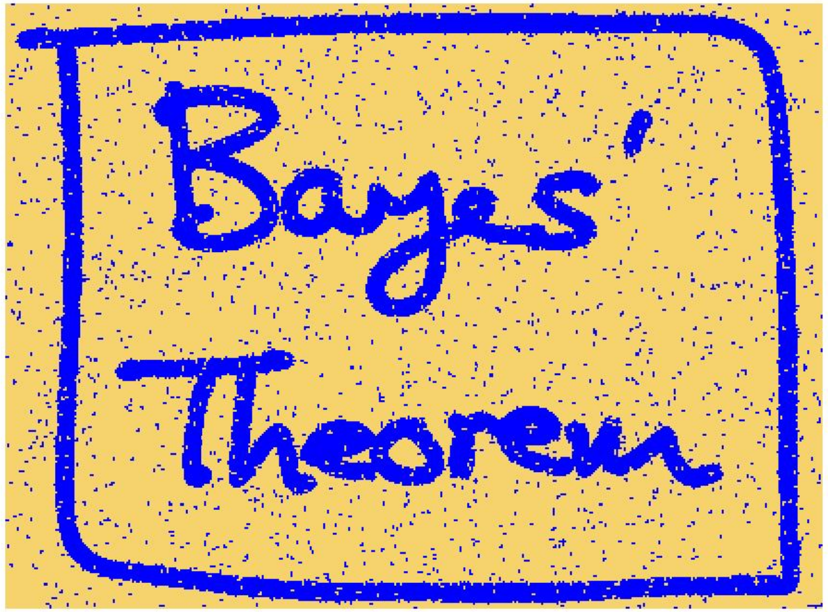


Noisy Image

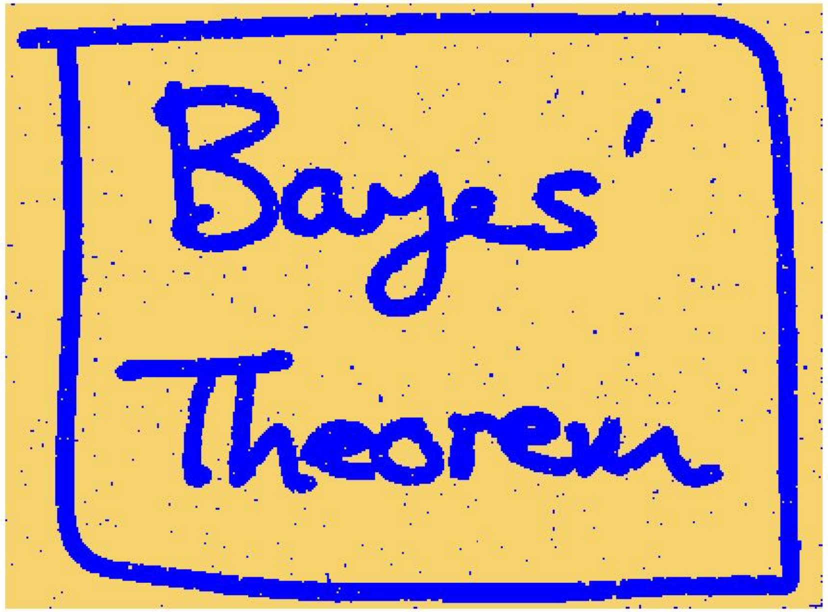


Restored Image (ICM)

Illustration: Image De-Noising



Restored Image (ICM)



Restored Image (Graph cuts)

常见的无向图模型

- 对数线性模型
  - 势能函数的一般定义为

$$\phi_c(\mathbf{x}_c|\theta_c) = \exp\left(\theta_c^{\mathrm{T}} f_c(\mathbf{x}_c)\right)$$

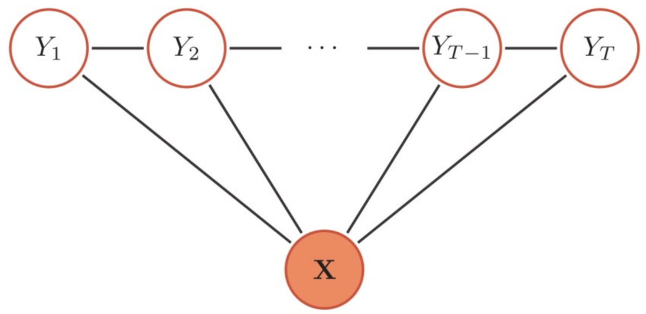
  - 联合概率p(x)的对数形式为

$$\log p(\mathbf{x}|\theta) = \sum_{c \in C} \theta_c^{\mathrm{T}} f_c(\mathbf{x}_c) - \log Z(\theta)$$

  - 也称为最大熵模型
- 条件随机场
  - y一般为随机向量
  - 条件概率p(y|x)

$$p(\mathbf{y}|\mathbf{x},\theta) = \frac{1}{Z(\mathbf{x},\theta)} \exp\left(\sum_{c \in \mathcal{C}} \theta_c^{\mathrm{T}} f_c(\mathbf{x},\mathbf{y}_c)\right)$$

(a) 最大熵模型



(b) 线性链的条件随机场

谢 谢

https://nndl.github.io/