计算机模式识别

-----概率图模型基础

主讲：图像处理与模式识别研究所
赵群飞

邮　　箱：zhaoqf@sjtu.edu.cn
办公室：电院 2-441
电　　话：13918191860

第4章　概率图模型基础

- 本章学习目标

✓明确判别式和生成式概率图模型的区别

✓掌握有向图模型的模型表示、条件独立性刻画，理解常见的有向图模型

✓掌握无向图模型的模型表示、条件独立性刻画，理解常见的无向图模型

✓掌握对树状结构因子图进行推理的和积算法和最大和算法

目录

概率图模型（PGM）简介

- 概率图模型是将概率论与图论相结合，以图的形式研究多元随机变量概率分布、推理和学习等问题的一类方法:
  - 其核心是以图的连接关系为基础，提供一种高效、可视化地表示随机变量之间条件独立性和联合概率分布的手段，比用一大堆概率公式更容易理解多个变量值的依赖关系；
  - 提供了模型中变量的推理方法，包括求解图中某些变量的边缘分布、以及变量使得联合后验分布最大的取值。

- 概率图模型是用概率论与图论对数据进行分析建模和推理的方法，从建模方式分为判别式模型和生成式模型：

- 判别式模型：对条件概率分布进行建模，直接构建从输入变量到输出变量之间的概率映射，可以直接用于预测。优点是直接刻画输入与输出的关系，避免对输入特征的建模，常见的模型有逻辑回归、条件随机场等。

- 生成式模型：对联合概率分布进行建模，即对输入特征和输出数据共同的概率分布建模，然后从联合分布计算后验分布再用于预测。好处是可以充分利用先验知识。常见的模型有朴素贝叶斯模型和隐马尔科夫模型。

概率图模型理论分为三部分内容：概率图模型的表示理论、推理方法和学习方法。

1. 概率图模型的表示理论可以分为结构表示和参数表示

- 结构表示是概率图模型的基础理论，以D-分隔、Hammersely-Clifford定理等为代表，揭示了联合分布的因子化表示和条件独立性（又称马尔可夫性）的等价性。

- 根据边的性质，概率图模型主要分为：
  - 有向无环图模型（即贝叶斯网络）
  - 无向图模型（即马尔可夫网络或马尔可夫随机场）
  - 混合图模型（同时包含有向边和无向边）

- 常见的无向图模型包括：
  - 条件随机场
  - 受限玻尔兹曼机
  - Ising模型等；
- 常见的有向无环图模型包括：
  - 隐马尔可夫模型
  - 混合高斯模型
  - 隐狄利克雷分配
- 常见的混合模型包括：
  - 深度置信网络等。

2. 概率图模型的推理方法分为精确算法和近似算法

- 经典的精确推理算法包括：变量消去法、信念传播算法、Junction Tree算法等。
- 近似算法分为两类：基于函数逼近的变分方法，如平均场算法，迭代信念传播算法；和基于随机采样的蒙特卡罗方法，如Importance Sampling，MCMC算法。
- 通常精确推理算法只用于链、树等简单图结构的问题中。由于计算复杂度的原因，求解实际应用问题时近似算法更加常用。

3. 概率图模型的学习可以分为**结构学习**和**参数学习**

- ✓ 图结构的学习已被证明一般是NP-Hard问题，尚无通用的学习算法；现有方法主要基于约束、搜索、动态规划、模型平均、混合策略等。
- ✓ 实际中通常的做法是针对具体问题人工设计图结构，例如：
  - 在混合高斯模型和话题模型中使用的混合加性结构
  - 在语音识别、手写字符串识别中使用的链式结构
  - 在图像降噪中使用的网格结构、层次化结构等

- 对于参数学习，极大化训练数据集上的似然函数是最常见的方式，但通常计算复杂度很高。
- 出于计算效率的考虑，经常使用其他目标函数对似然函数进行近似，如：
  - 似然函数的变分下界（Evidence Lower Bound）
  - 分段似然（Piecewise Likelihood）
  - 伪似然（Pseudo Likelihood）
  - 评分匹配（Score Matching）
  - 矩匹配（Moment-Matching）
  - 对抗训练（Adversarial Training）

目录

4.1 有向图模型

有向模型图（Directed Graphical Model），也叫贝叶斯网络（Bayesian Network），通常使用有向无环图表示概率分布。

- 其特点是即可以表示所有变量的联合分布，又可以表示直接相连的节点之间的条件分布。
- 相关的的节点用有向边连接，表示节点间的依赖关系或者说因果关系，因此也有称之为因果网络。

4.1.1 模型表示

一个贝叶斯网络由两部分组成：

1）有向无环图 $\mathcal{G}=(V,\mathrm{E})$，其中 V 表示有向图中节点的集合，E 表示图中有向边的集合。

2）父节点到子节点的条件概率分布。

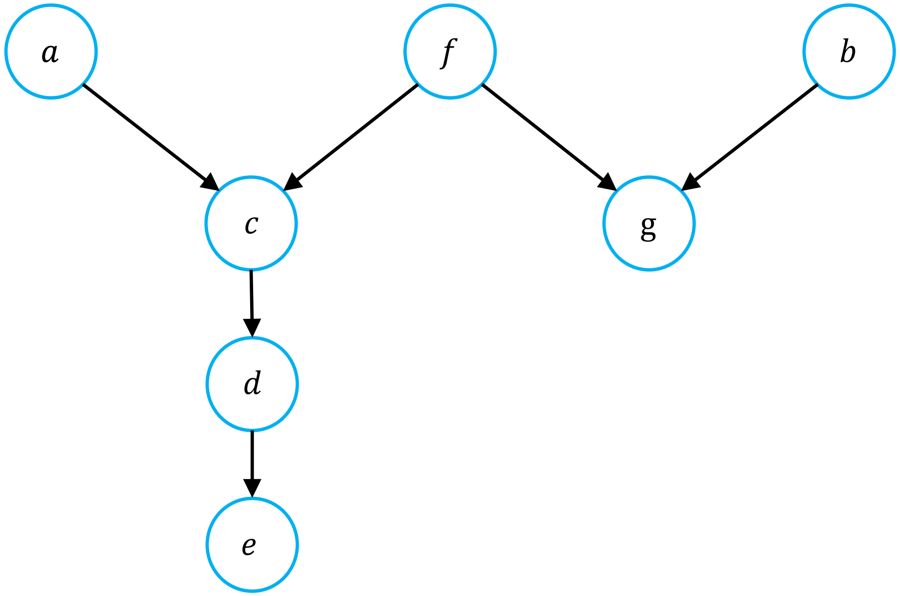
如果 $\forall a\in V,\forall b\in V$， $a\rightarrow b\in E$，

那么 a 称为 b 的父节点，

则 b 称为 a 的子节点。

每个节点可有多个父节点

也可以有多个子节点



- 用贝叶斯网络可描述变量之间的依赖关系，同时也刻画了变量之间的独立关系。
- 通过提供因子化形式的联合分布，等价地暗含了所有变量中存在的条件独立性
- 考虑网络中有K个随机变量的联合分布。根据概率分布的乘法运算法则，K个随机变量的联合分布为：

$$p(x_1,x_2,...,x_K)=p(x_1)p(x_2|x_1)\cdots p(x_K|x_1,x_2,...,x_{K-1})$$

- 如果每个节点对应一个随机变量，这个联合分布可以表示为有K个节点的有向图，有向边对应于上式右边的每个条件分布。形成一个有向无环图。

➢ 贝叶斯网络定义的紧凑的联合分布表示为：

$$p(x_1, x_2, \ldots, x_K) = \prod_{k=1}^{K} p\left(x_k \mid \mathrm{Pa}_{x_k}\right)$$

其中，Pa_{x_k}表示节点x_k的父节点。

- 上式给出了贝叶斯网络的联合分布，这个因子化形式暗含了模式中的条件独立性。

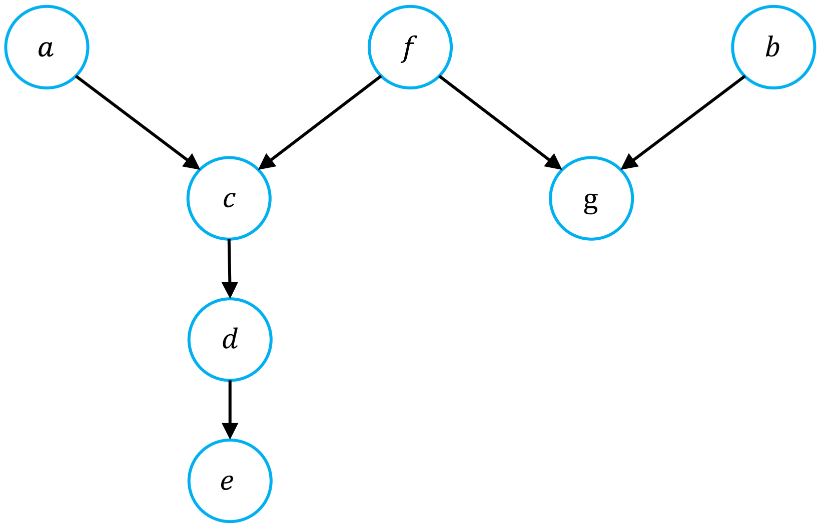


图4-2 贝叶斯网络示例

根据贝叶斯网络的紧凑的联合分布，可得上图贝叶斯网络的联合分布为

$$p(a, b, \ldots, g) = p(a)\,p(b)\,p(f)\,p(c \mid a, f)\,p(g \mid b, f)\,p(d \mid c)\,p(e \mid d).$$

4.1.2 条件独立性

独立：设两个随机变量a和b，如果它们的联合概率分布满足：

$$p(a,b) = p(a)p(b)$$

则称随机变量a和b是相互独立的，记为 $a \perp b$.

条件独立：如果在给定一个额外的随机变量c后，a和b的联合分布可以写作：

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

那么，称a和b是条件独立的，记为 $a \perp b|c$.

局部条件独立：贝叶斯网络中每一个节点在给定其父节点的条件下与其他非后代节点条件独立。

由一个有向无环图 $\mathcal{G}$ 及其联合分布$p(x_1, x_2, \cdots, x_K)$构成的一个贝叶斯网络，其联合分布满足条件独立性：

$$\forall k, x_k \perp \mathrm{NonDesc}(x_k) | \mathrm{Pa}(x_k),$$

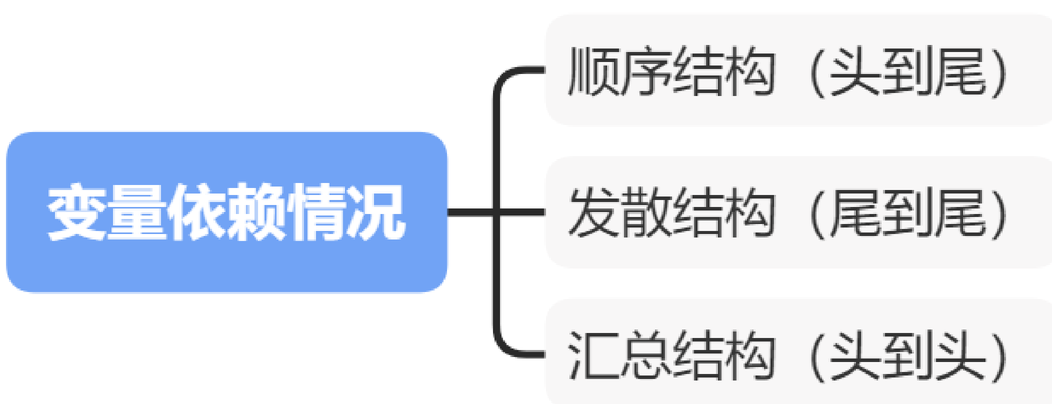
其中，$\mathrm{NonDesc}(x_k)$表示除$\mathrm{Pa}(x_k)$之外x_k的非后代节点。

- 联合分布的紧凑表示与其局部条件独立性是对贝叶斯网络等价的描述（见p66证明）

◆ 通过联合分布的紧凑表示或者通过局部条件独立性的形式化语义，都可以分析出贝叶斯网络中变量的一些条件独立性，但是这两种方法并没有将所有的独立性情况包括。

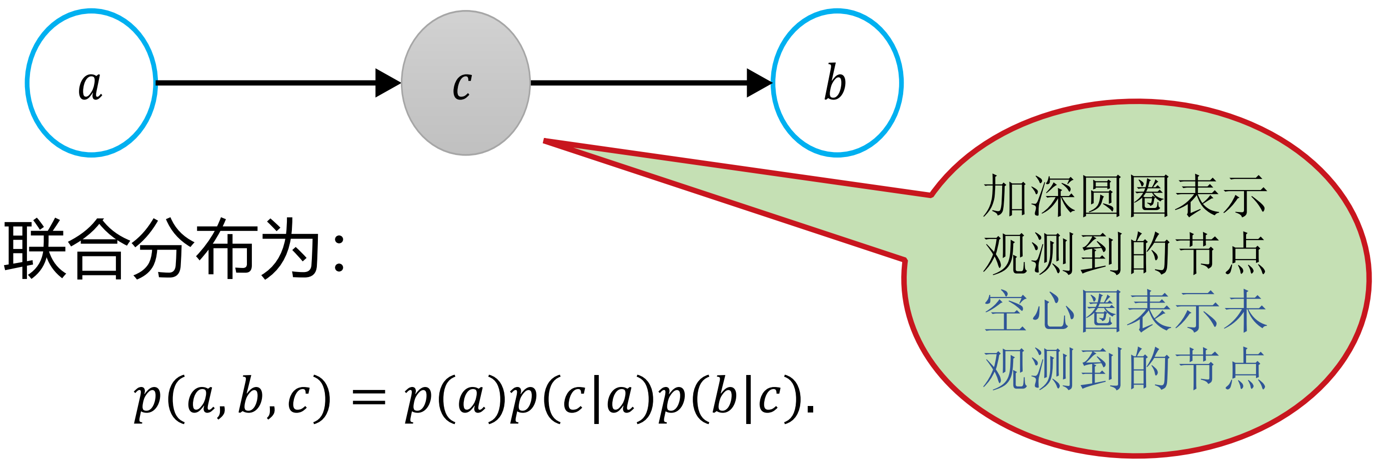
◆ 事实上，通过图的一些特殊结构和规则可以简单直观地得到所关心变量的条件独立性。

一般地，三种基本的变量依赖情况，对应三种不同的图结构：



顺序结构：节点c连接了一个箭头的头部和另一个箭头的尾部。

顺序结构具有条件独立性：在给定c的条件下，a和b条件独立。



概率图模型的联合分布为：

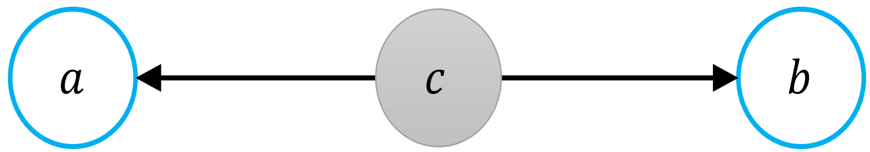
$$p(a,b,c)=p(a)p(c|a)p(b|c).$$

因此

$$p(a,b|c)=\frac{p(a,b,c)}{p(c)}=\frac{\overbrace{p(a)p(c|a)}^{p(a|c)p(c)}p(b|c)}{p(c)}=p(a|c)p(b|c).$$

发散结构：节点c连接两个箭头的尾部。

发散结构具有条件独立性：在给定c的条件下，a和b条件独立。



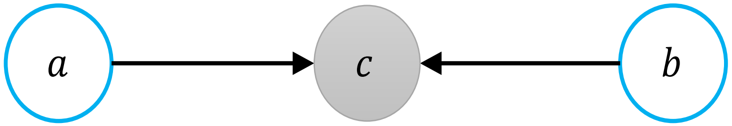
概率图模型的联合分布为：

$$p(a,b,c) = \; p(a|c)p(b|c)p(c).$$

$$p(a,b|c) = \frac{p(a,b,c)}{p(c)} = p(a|c)p(b|c).$$

汇总结构：节点c连接了两个箭头的头部。

汇总结构不具有条件独立性：在给定c的条件下，a和b条件不独立。



概率图模型的联合分布为：

$$p(a,b,c) = p(a)p(b)p(c|a,b).$$

关于c求积分或者求和后得到a和b的联合分布可以表示为各自边缘分布的乘积：

$$p(a,b) = p(a)p(b).$$

d-分隔规则 *(刻画更复杂的条件独立性)*

设 $\mathcal{G} = (V, E)$ 是一个贝叶斯网络（有向图模型），集合A, B, C是V中相互不相交的子集，其中C中的节点是被观测到的。考虑从A中节点到B中节点所有可能的路径， 如果路径上存在一个节点满足如下两个条件之一，那么该条路径是被阻隔的。

- 该节点是被观测的（在集合C中），并且它与所连接的节点具有发散结构（尾到尾）或者顺序结构（头到尾）；
- 该节点及其后代节点都未被观测（不在集合C中），并且它与所连接的节点具有汇总结构（头到头）。

如果从$\boldsymbol{A}$中节点到B中节点的所有路径都被阻隔，就可以说从A到B的路径是被阻隔的，或者叫d-分隔的。

对于一个贝叶斯网络，在给定观测的节点集合C的条件下，如果A到B的所有路径都被阻隔，就可以说A和B在给定C的情况下条件独立，图中所有变量的联合概率分布将会满足$\boldsymbol{A} \perp \boldsymbol{B} | \boldsymbol{C}$。

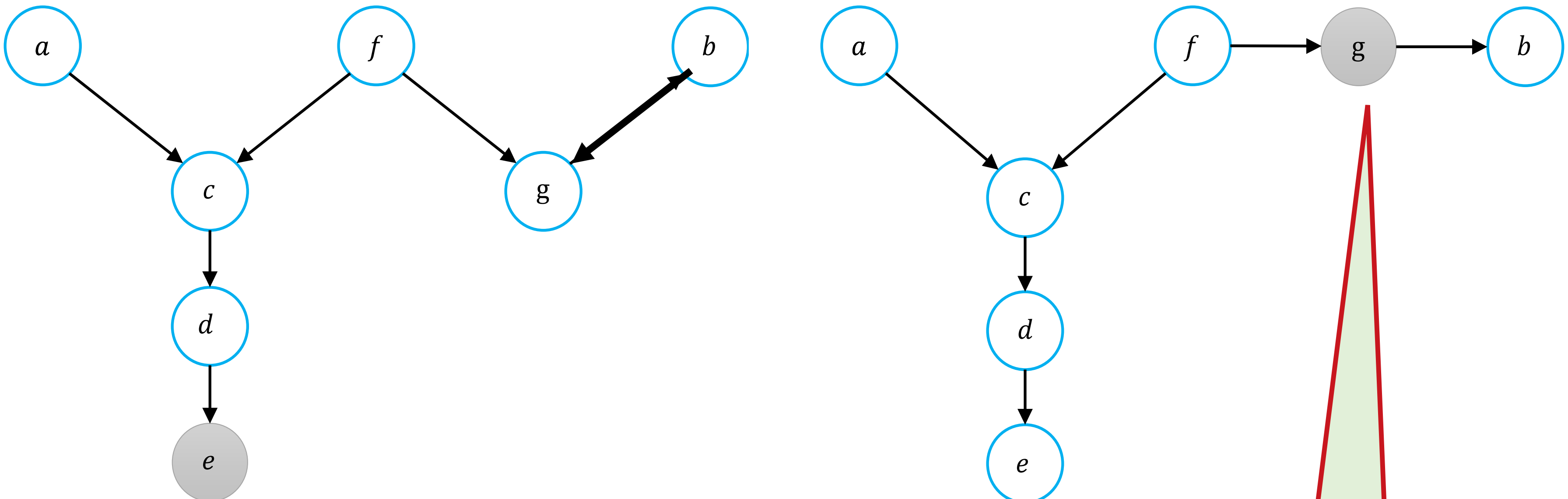


图4-7 使用d-分隔判断条件独立性

给定节点e后，a和b不是条件独立的

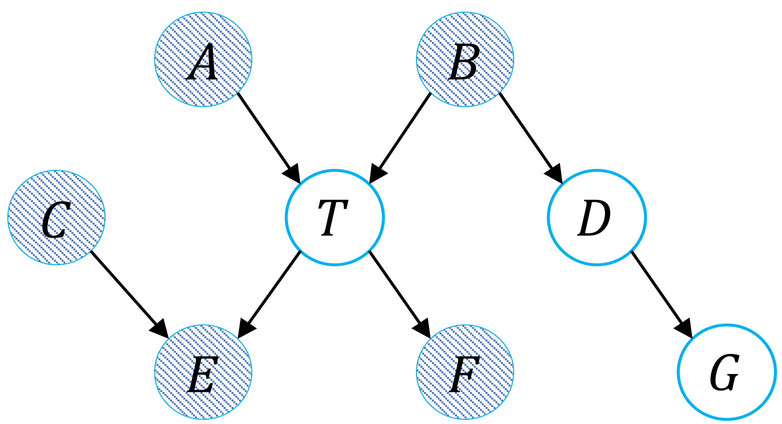
给定节点g后，a和b是d-分隔的$a \perp b|g$

马尔可夫毯（Markov blanket)

在所考虑的随机变量的全集U中，对于给定的变量$x \in U$和变量集$MB \subset U$，且$x \notin MB$)，若有

$$x \perp \{U - MB - \{x\}\} | MB,$$

则称满足上述条件的最小变量集MB为x的马尔可夫毯。



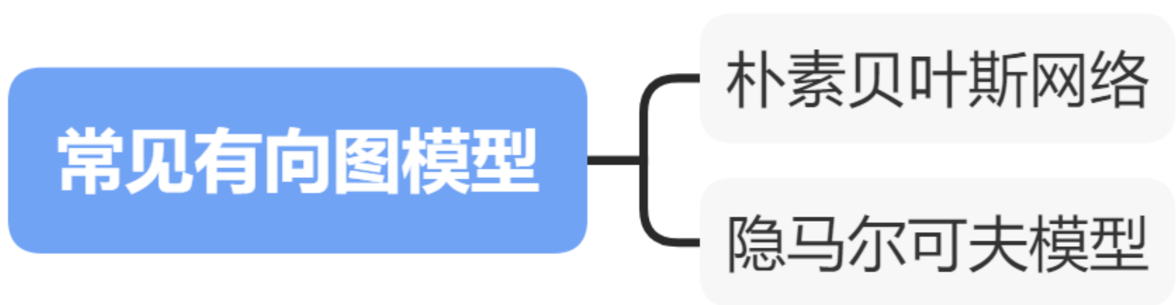
一个节点的马毯包括该节点的父节点、子节点以及子节点的共同父节点

有向图的马尔可夫毯示意图

(图中纹理填充的节点构成了节点T的马尔可夫毯)

4.1.3 常见的有向图模型

- 有向图模型是一类模型框架，针对不同的应用，研究者们提出了很多不同的模型实例，例如，用于分类的朴素贝叶斯和用于时序数据建模的隐马尔可夫模型等。
- 本节仅以朴素贝叶斯和隐马尔可夫模型为例，介绍各自的模型假设及其概率图模型表示。



朴素贝叶斯网络（naïve Bayes）

假设$\mathbf{x}$是一个D维的样本向量，其中每一个元素x_d表示$\mathbf{x}$的一个特征或属性，$z = 1,2,\cdots,C$ 表示数据$\mathbf{x}$所属的类别，例如$z = c$ 表示数据属于第c类。在朴素贝叶斯分类器中，D个特征在已知类别的条件下相互独立。

如图，其中观测到的特征变量$x_1, x_2, \cdots, x_D$依赖于类别变量z，并且在已知类别z的条件下，特征之间$x_1, x_2, \cdots, x_D$是条件独立的。

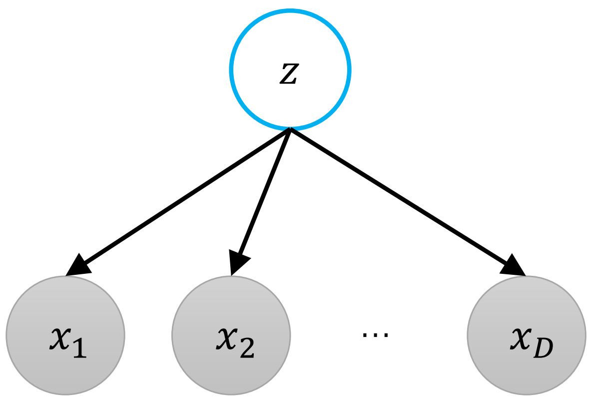


图4-9 朴素贝叶斯网络示意图

由于朴素贝叶斯网络中各个特征之间是条件独立的，因此每一个样本的类条件概率分布有如下表示：

$$p(\mathrm{x}|z=c)=\prod_{d=1}^{D}p(x_d|z=c)$$

已知类条件概率分布，贝叶斯分类器通过计算后验概率 $p(z=c|\mathbf{x})$进行决策。根据贝叶斯公式可得到后验概率为：

$$p(z=c|\mathrm{x})=\frac{p(z=c)p(\mathrm{x}|z=c)}{p(\mathrm{x})}=\frac{p(z=c)}{p(\mathrm{x})}\prod_{d=1}^{D}p(x_d|z=c).$$

由于对于所有类别$p(\mathbf{x})$都是相同的，朴素贝叶斯分类器等价于寻找一个c类别，使得$p(z=c)\prod_{d=1}^{D}p(x_d|z=c)$最大。

隐马尔可夫模型（hidden Markov model）

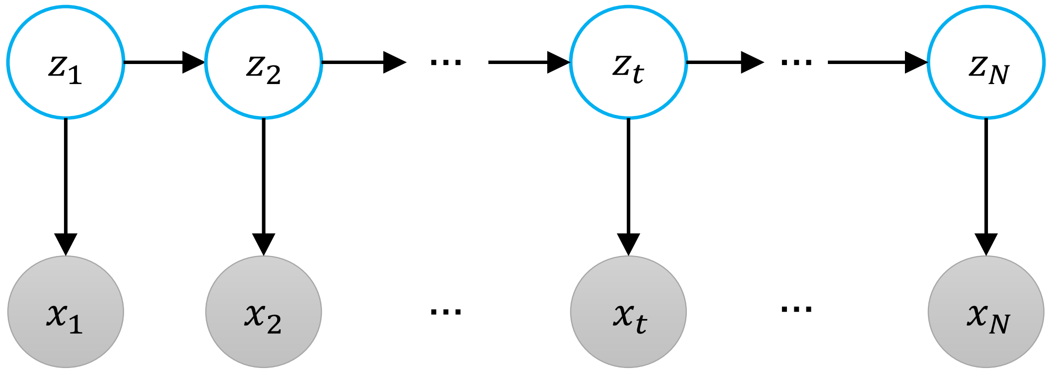


图4-10 隐马尔可夫模型示意图

在任意时刻t，给定z_t，观测变量x_t的取值与其他隐变量无关。同时，对于任意时刻的隐变量z_t，如果给定其前一时刻的隐变量z_{t-1}，则与更早时刻的隐变量没有关系。

图4-10对应的联合概率分布为：

$$p(x_1, z_1, x_2, z_2, \ldots, x_N, z_N) = p(z_1)p(x_1|z_1)\prod_{t=2}^{N} p(z_t|z_{t-1})p(x_t|z_t).$$

目录

4.2 无向图模型

- 无向图模型，也称为马尔可夫随机场（Markov Random Field）或马尔可夫网络。
- 可描述一组具有局部马尔可夫性的随机变量的联合概率分布模型
- 节点之间的边不带方向，表示变量之间的相互约束或兼容关系
- 相连节点的关系通常使用势能函数来刻画，而不是条件概率分布。

4.2.1 模型表示

- 对于一个随机变量的集合

$X = \{x_1, x_2, \cdots, x_K\}$

和相应的有K个节点的无向图 $\mathcal{G}(V, E)$（可以有循环），图中的节点表示随机变量。

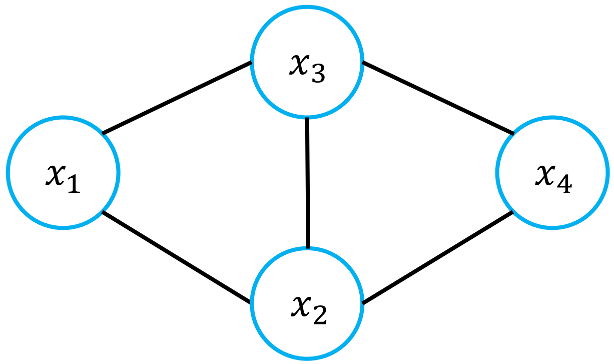


图4-11 无向图模型示例

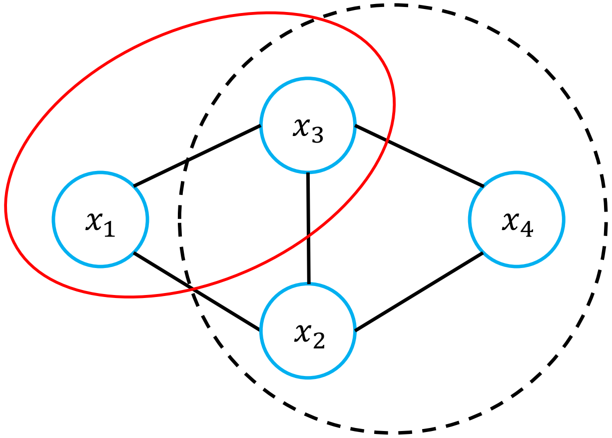
- 如果$\mathcal{G}$满足局部马尔可夫性质，即任一变量x_k在给定它的邻居的情况下条件独立于所有其他变量，表示为x_k在给定邻居变量和给定其他所有变量条件下的概率分布相同，即：

$$p(x_k | X_{\backslash k}) = p(x_k | X_{\mathrm{ne}(k)}),$$

那么 $\mathcal{G}$ 就构成了一个**马尔可夫随机场**。

式中$X_{ne(k)}$表示变量x_k的邻居的集合，$X_{\backslash \mathrm{k}}$表示除$x_k$之外其他变量的集合。

- 无向图模型的联合分布一般以全连通子图为单位进行分解。
- 无向图中的一个全连通子图，称为团（clique），即团内的所有节点之间都有边相连。
- 如果一个团不被其他团包含则称之为最大团



- 图4中除单点团外共有7个团，包括 $\{x_1,x_2\}$、$\{x_1,x_3\}$、$\{x_2,x_3\}$、$\{x_3,x_4\}$、$\{x_2,x_4\}$、$\{x_1,x_2,x_3\}$和$\{x_2,x_3,x_4\}$
- 实线中的节点和边构成一个团，虚线中的节点和边构成一个最大团。

无向图中的联合概率分布可以分解为一系列定义在最大团上的非负函数的乘积形式，即：

$$p(X) = \frac{1}{Z} \prod_{c \in \mathcal{C}} \phi_c\left(X_c\right),$$

其中$\mathcal{C}$为$\mathcal{G}$中的最大团集合，$\phi_c(X_c) \geq 0$是定义在团c上的势能函数（potential function），Z是配分函数（partition function），用于将乘积归一化为概率分布形式，即：

$$Z = \sum_{X \in \mathcal{X}} \prod_{c \in \mathcal{C}} \phi_c\left(X_c\right),$$

其中$\mathcal{X}$为随机变量的取值空间。

4.2.2 条件独立性

无向图中的局部马尔可夫性可以表示为：

$$\boldsymbol{x_k} \perp \boldsymbol{X_{\backslash ne(k),\backslash k}} | \boldsymbol{X_{ne(k)}}$$

其中$X_{\backslash \mathrm{ne}(k),\backslash k}$表示除$X_{\mathrm{ne}(k)}$和$x_k$之外的其他变量。

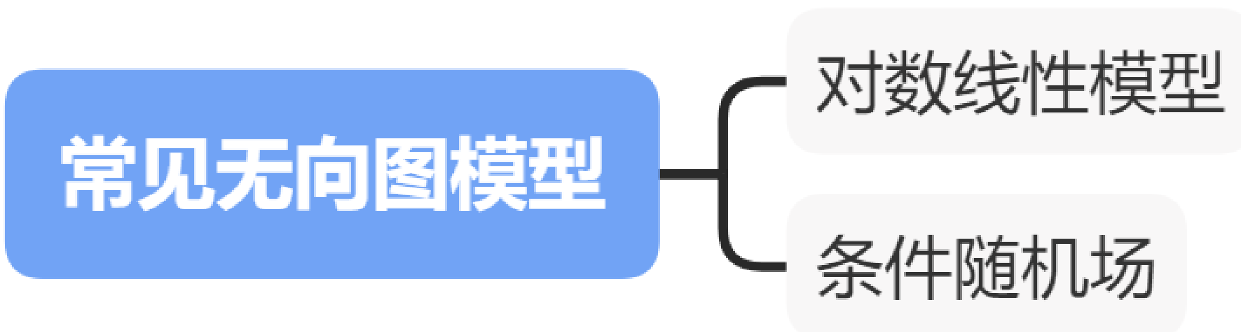
- 局部马尔可夫性和与联合概率分布表示是等价的

Hammersley-Clifford定理：如果一个分布$p(X) > 0$满足无向图中的局部马尔可夫性，即$x_k \perp X_{\backslash \mathrm{ne}(k),\backslash k} | X_{\mathrm{ne}(k)}$，当且仅当$p(X)$可以表示为一系列定义在最大团上的非负函数的乘积形式，即

$$p(X) = \frac{1}{Z} \prod_{c \in \mathcal{C}} \phi_c\,(X_c),$$

4.2.3 常见的无向图模型

很多机器学习模型可以使用无向图模型来描述，比如对数线性模型（也叫最大熵模型）和条件随机场等。



本节以对数线性模型，介绍其模型假设及其概率图模型表示。

对数线性模型

➢ 势能函数一般定义为：

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

其中　函数$f_c(\mathbf{x}_c)$为定义在$\mathbf{x}_c$上的特征向量，

参数θ_c为权重向量。

♥ 这样联合概率分布的对数形式为：

$$\ln p\,(\mathbf{x}|\theta) = \sum_{c\in C} \theta_c^\top\, f_c(\mathbf{x}_c) - \ln Z\,(\theta)$$

其中$\boldsymbol{\theta}$代表所有势能函数中的参数$\{\theta_c\}$。

♥ 这种形式的无向图模型也称为对数线性模型或最大熵模型。

如果用对数线性模型来建模条件概率分布 $p(\mathbf{y}|\mathbf{x})$,

那么带有参数的条件概率分布表示 $p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$为:

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

其中

$$Z(\mathbf{x},\theta)=\sum_{\mathbf{y}}\exp(\theta^{\top}f(\mathbf{x},\mathbf{y}))$$

目录

4.3 图模型中的推理

- 图模型的推理是计算图中某些变量的边缘分布，以及求解使得联合概率分布最大的某些变量的值
- 树状的概率图模型，可以统一表现为树状结构因子图，其推理算法可采用和积算法或最大和算法。

4.3.1 链式结构

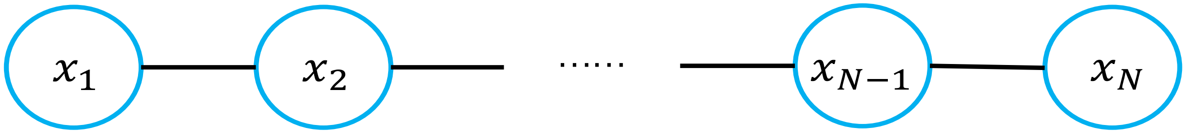


图4-16 链式无向图示例

其联合分布为：

$$p(\mathrm{x}) = \frac{1}{Z}\psi_{1,2}(x_1, x_2)\psi_{2,3}(x_2, x_3) \ldots \psi_{N-1,N}(x_{N-1}, x_N).$$

由于没有观测节点，节点x_n对应的边缘概率分布为

$$p(x_n) = \sum_{x_1} \ldots \sum_{x_{n-1}} \sum_{x_{n+1}} \ldots \sum_{x_N} p\,(\mathrm{x}).$$

节点x_n的边缘概率分布可以分解为两个因子的乘积x归一化系数

$$p(x_n) = \frac{1}{Z}\mu_\alpha(x_n)\mu_\beta(x_n),$$

其中， $\mu_\alpha(x_n)$定义为从节点x_{n-1}到节点x_n的沿着链向前传递的消息，

$\mu_\beta(x_n)$定义为从节点x_{n+1}到节点x_n的沿着链向后传递的消息。

$$\mu_\alpha(x_n) = \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n)\left[\sum_{x_{n-2}} \cdots\right] = \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n)\mu_\alpha(x_{n-1}),$$

$$\mu_\beta(x_n) = \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})\left[\sum_{x_{n+2}} \cdots\right] = \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})\mu_\beta(x_{n+1}).$$

4.3.2 树状图

♥ 树状的概率图模型

- 任意一堆节点之间只有一条路径，不存在环路，
- 有一个没有父节点的节点，称为根；其他节点只有一个父节点。

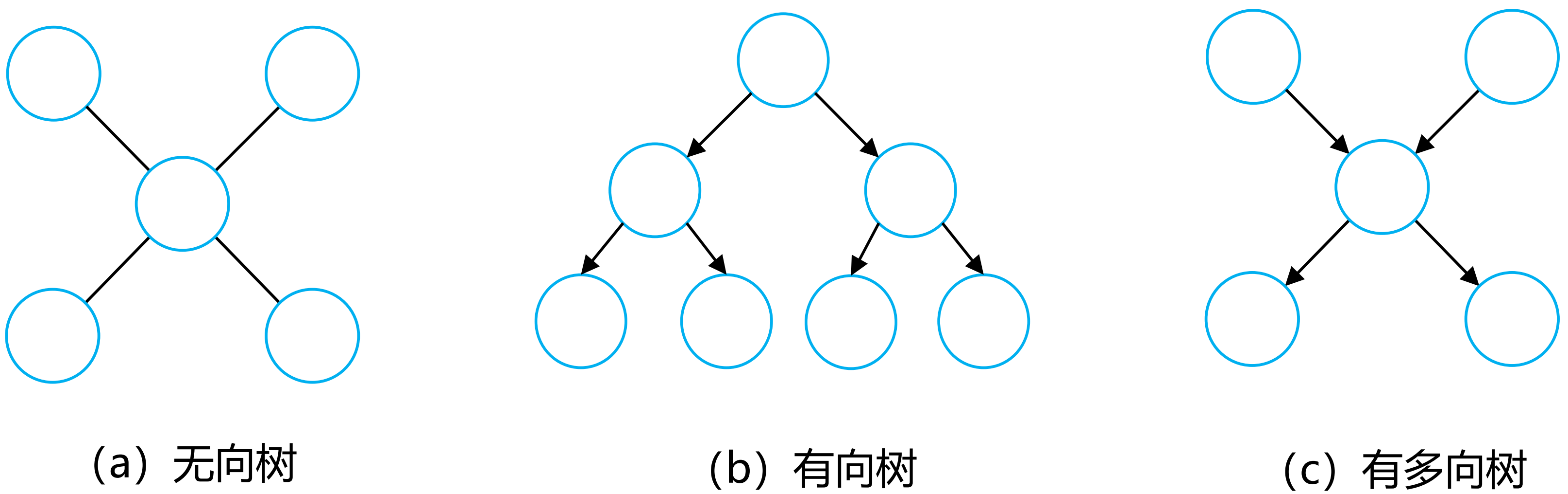


图4-17三种树图的概率图模型示意图

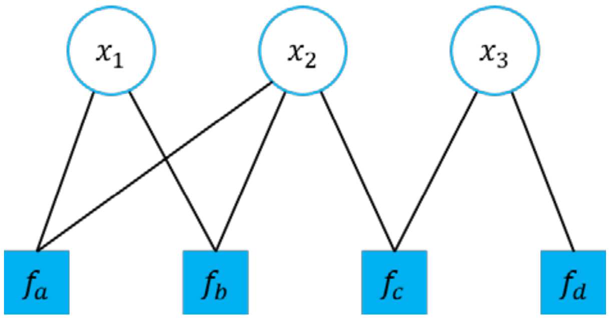
4.3.3 因子图

因子图与图模型的对应关系：

- 因子图中的变量节点与对应图模型中的变量节点相同；
- 因子图中对应图模型中同一因子的变量节点之间存在一个因子节点；
- 因子图中的边都是无向边，连接因子节点与相对应的变量节点。

➢ 概率图模型中随机变量的联合概率分布$p(X)$可以写成因子图中各个因子$f_s(x_s)$的乘积形式，即：

$$p(X)=\prod_s f_s(x_s).$$



➢ 图中变量的联合概率分布可以表示为：

$$p(X)=f_a(x_1,x_2)f_b(x_1,x_2)f_c(x_2,x_3)f_d(x_3).$$

4.3.4 和积算法

和积算法可以利用因子图的结构实现如下两个目标：

① 获得对因子图中节点的边缘分布的高效确切推理；

② 在需要计算多个边缘分布时实现计算共享。

和积算法的核心数学思想是交换求和运算和求积运算的次序。

交换求和与求积运算的依据是乘法分配律：

$$\sum_{i,j,k} x_i\, y_j z_k = (\sum_i x_i) \times (\sum_j y_j) \times (\sum_k z_k).$$

和积算法体现了消息传递的思想。消息分为两种：

① 从因子节点到变量节点的消息；

② 从变量节点到因子节点的消息。

- 图模型：图模型是由图结构构成的，其中节点表示随机变量，边表示变量之间的依赖关系。
- 贝叶斯网络：是有向图模型，每个节点都有一个相关的条件概率分布。
- 马尔可夫网络：是无向图模型，每个团都有一个相关的势函数。
- 条件独立：根据图中节点的连接方式，我们可以写出这种形式的条件独立陈述：「给定 Z，则 X 与 Y 相互独立」。
- 参数估计：根据给定的一些数据和图结构来填充 CPD 表或计算势函数。
- 推理：给定一个图模型，我们希望解答有关未被观察的变量的问题，这些问题通常属于以下问题范围：边际推理、后验推理和 MAP 推理。

在一般图模型上的推理的计算非常困难。我们可以将推理算法分成两大类——精准推理和近似推理。无环图中的变量消除和置信度传播是精准推理算法的例子。近似推理算法对大规模图而言是必需的，而且通常属于基于采样的方法或变分法。

贝叶斯网络模型学习 （图模型的学习）

贝叶斯网络模型模型的参数学习，一般会假设结构一致，然后从样本数据中学习每个变量的概率分布。概率分布的形式一般会预先设定，如多项式分布、高斯分布和泊松分布等。贝叶斯网络模型的结构学习是贝叶斯网络学习的最主要的部分，此时贝叶斯网络的参数和结构都未知，需要从样本数据中找到与数据匹配度最好的网络结构。当确定网络结构之后，参数学习知识相对简单的参数估计问题。结构学习的算法主要分为三类：基于约束(Constraint based, CB) 的学习算法、基于评分搜索(Scoring and searching, SS) 的学习算法和混合学习算法。

马尔可夫随机场模型学习（图模型的学习）

马尔可夫随机场模型的学习任务比贝叶斯网络模型的学习更加复杂。在CRF模型的参数学习中，其配分函数耦合了结构中的所有参数，使得参数学习问题不能分解，不能分别独立估计每个局部参数。而且马尔可夫随机场的最优参数没有解析解，所以一般使用迭代方法求解最优参数，如梯度下降法。庆幸的是，迭代中的似然目标函数为凹函数，迭代方法能保证收敛至全局最优解。但是迭代中的每一步都需要在网络中进行推理，使得计算成本相当高。MN 的结构学习也需要计算划分函数，所以其参数学习存在的问题对结构学习有很大的影响。

CMU Eric Xing教授每年开的10708 PGM课程
10708 Probabilistic Graphical Models。

机器学习理论与实战（十四）概率图模型02
https://blog.csdn.net/marvin521/article/details/10692153

机器学习理论与实战（十五）概率图模型03
https://blog.csdn.net/marvin521/article/details/10858655

Coursera概率图模型（Probabilistic Graphical Models）第二周编程作业分析

一文读懂机器学习概率图模型（附示例和学习资源）
https://cloud.tencent.com/developer/article/1032273

谢 谢！