

概率图模型中的变分近似推理方法

刘楠

(吉林大学材料科学与工程学院, 吉林 长春 130000)

摘 要: 概率图模型中的变分近似推理方法是将概率图模型中的图论和概率论加以结合, 分析多个变量之间的依赖关系, 用计算机生物学中的语言处理、视觉处理的应用, 采用边缘概率和概率状态推理出概率图模型的核心问题。本文在变分近似推理的框架下, 对于变分推理的研究成果, 即概率图模型中的变分近似推理方法进行系统的归纳和思路梳理, 对于目前较多使用的近似推理方法进行综述, 归纳了近似算法的单调性以及收敛性的性质, 展望了未来概率图模型近似推理方法的研究方向。

关键词: 概率图模型; 贝叶斯网; 变分法

概率图模型由复杂的依赖关系的多个变量提供统一的框架, 形成了复杂的多元统计关系, 内涵图论和概率, 特点是紧凑而简洁。从 20 世纪 90 年代被加以关注后, 概率图的模型推理形成了自己独特的理论价值, 在学习和推理问题上进行了几十年的研究, 得到了一定的研究成果。采用计算机视觉和语言处理的方法, 对于概率图的模型加以生物学的计算, 应用于控制理论、统计物理学领域中。形成了一系列的应用, 包括立体分割、图像匹配、图像合成以及纹理的去噪等, 将概率图模型的推理问题、信号处理问题等加以标注, 形成了通信信号处理的解码等算法, 经过概率图的模型的新都传播推理校验, KALMAN 滤波算法、TURBO 解码算法等, 最终将生物信息学的问题归结到概率图模型的推理问题^[1]。

进行概率图的梁旭结构的归结和组合, 采用本身的近似推理方法, 包含了的种类为种, 一种是变分法, 一种是采样法, 前者是将目标函数进行最小化, 采用优化的方法进行问题的解决。后者是从概率图的样本近似边缘加以联合部分, 然后进行最小概率状态的推理。两种方法都是较为广泛使用的推理方法。经过对偶优化的角度加以推理后, 得到了目标范围、约束方法、优化函数三个方面的推理的算法。集中算法种, 分解的思想是变分近似的灵魂, 近似的求解上, 边缘的概率问题的分析采用的是用简单的方法代替反复的算法。使用局部变量的边缘概率的组合来进行多变量联合概率的计算。在对概率状态问题进行近似求解的过程中, 分解思想的表现方法就是将复杂的问题简单化, 拆分成容易解决的问题, 然后把每个子问题加以整合, 得到原始问题的近似解。

1 概率图画的模型和推理

1.1 概率图的模型是一种有效的模型, 采用 G、V 等加以表示, 包含了结点结合状态分析, 离散连续空间标记, 空间概率图模型的模型等, 通过对模型的变分近似推理, 得到了较为广泛的模型的形态。

采用贝叶斯网进行概率图模型和马尔科夫随机场的计算, 显示出两个变量之间的依赖关系。通过条件概率表和势函数来表示依赖关系, 得到了公式进行联合概率分布的计算:

$$p(\mathbf{x}) = \prod_{s \in V} p_s(\mathbf{x}_s | \mathbf{x}_{\pi(s)})$$

其中, \mathbf{x}_s 为变量 x_s 的父结点的集合, 将模型进行等价地转化, 得到了无向图模型, 等价的意思就是将联合概率以及变量之间进行分布, 采用独立关系不变的方式, 在实际应用上, 扩大马尔科夫随机场和贝叶斯网的表示范围, 贝叶斯网的信息量, 包含了有向变的因果关系, 在马尔科夫随机场中无法表示。

使用因子图的表示方法对贝叶斯网和马尔科夫的随机场进行计算, 显示了一个 2×2 的因子图, 加以区域的概率图模型的表示, 利用因子分解的方法, 采用多个变量联合分布的方法, 将独立性的假设紧凑地加以表示, 其变量可以在图结构中独立, 其联合概率分布可以采用局部函数的乘积加以表示^[2]。

1.2 概率图模型的应用需要解决最大概率状态等问题, 原因之一是概率图模型应用的最终目的是得到需要的推理结果, 最大概率状态计算边缘化在很多结构学习和参数学习算法中是求取部分变量, 通过样本数据估计出最符合数据间依赖关系的概率图模型结构, 在给定概率图模型结构的情况下, 得到所有变量的边缘概率或者最大概率状态。参数学习和概率推理三个问题结构学习, 如条件概率分布等, 需要概率推理算法作为中间步骤, 解决概率推理问题是概率图模

型研究及应用的重要问题。在给定样本数据的情况下进行概率推理, 将概率推理所关心的问题——参数学习加以估计, 得出各个变量之间定量的依赖关系以及概率推理问题包括计算部分变量的边缘概率。

2 精确推理方法

探索各种条件下的精确推理方法、精确推理的研究思路是当前研究工作人员孜孜以求的目标, 对特殊图结构进行精确推理的概率图模型的构建, 得到特殊参数条件下的多项式级复杂度的精确推理方法, 形成概率图模型上不可精确计算的求解结果, 包括(团)树状概率图、基于求最大的边缘概率以及单环概率图、基于求和的边缘概率变量状态为二值的概率图模型。

2.1 边缘概率和联合概率的关系是边缘概率是联合概率的边缘化, 基于求和的边缘概率精确推理树状图模型上的基础和边缘概率 Sum-marginal, 它们之间存在两种关系: 概率图模型上边缘概率和联合概率之间的关系, 联合概率可以用边缘概率的组合来表示。定义为:

$$p_s(\mathbf{x}_s) = \sum_{\mathbf{x}_{\setminus \mathbf{x}_s}} p(\mathbf{x})$$

2.2 概率图模型所研究的自由能实质是一种距离的度量对于非树状概率图模型, 团树上的信度传播算法的复杂度是团树最大团中变量数目的指数级, 直接运用信度传播算法得到精确解, 然后利用团树上进行信度传播算法, 得到原始图模型的精确推理结果, 把非树状图模型化为团树(又称为邻接树)模型^[3]。

3 变分近似推理的发展阶段

变分法的思路是首先解决优化问题的环节将推理问题转化为优化问题, 然后利用原始推理的近似解, 近似方法的带优化, 变分近似推理理论大致分为三个研究阶段来解决推理问题。不同阶段的研究思路的不同: 信度传播阶段、自由能优化阶段和对偶优化阶段。

基于变分法的边缘概率采用“分解”的思路, 对边缘概率的组合进行分布, 得到近似所有变量的联合概率。这种“分解”思路的具体表现为: 基于分解的边缘对自由能的分解、对概率的近似求解、对联合熵的分解、对联合概率的分解等^[4]。

绝大部分概率图模型的推理问题无法精确求解, 概率推理表现出简约而紧凑的表示能力, 对于概率图模型研究及应用的核心问题, 是解决这些概率推理问题的唯一途径, 这一点在很多实际问题上都得到了验证。在具体应用方面, 如计算机视觉中的跟踪和行为识别问题、各种益智游戏的推理问题等实际问题、机器信息学领域的数据挖掘、交通网络、电网优化动态调度、生物优化结构等方面, 都表现出巨大的应用潜力。

参考文献

- [1] 刘建伟, 崔立鹏, 黎海恩, 等. 概率图模型推理方法的研究进展[J]. 计算机科学, 2015, 42(4): 1-18, 30.
- [2] 黎海恩, 刘建伟, 罗雄麟, 等. 概率图模型的变分近似推理[C]. // 2013 年中国智能自动化会议论文集. 2013: 6-11.
- [3] 龚声蓉, 叶芸, 刘纯平, 等. 基于在线消息传递的主题追踪方法[J]. 计算机学报, 2015, 38(2): 249-260.
- [4] 刘建伟, 崔立鹏, 黎海恩, 等. 概率图模型推理方法的研究进展[C]. // 2014 湖北省计算机学会学术年会论文集, 2014: 1-18, 30.