小数据集下基于修正乘性协同约束的 BN 参数学习

任智芳，陈海洋*，环晓敏，尚珊珊
（西安工程大学电子信息学院，西安，710048）

摘要　在一些特定情况下，获取充足样本十分困难，导致最大似然估计算法学习到的 BN 参数精度往往较低，并且一些实际应用领域中已涉及多父节点协同影响约束的问题。对此，通过借鉴 PAVA 保序回归算法思想，提出了一种小数据集下基于修正乘性协同约束的 BN 参数学习方法。首先，判断已知样本数据中多父节点部分的参数是否满足乘性协同约束；其次，把不满足乘性协同约束的左右两边划分为整体，用 PAVA 算法分别对其进行调整，针对调整后的整体，根据不同父节点组合状态对应的样本数据量，给出 3 种权值不同的校正方法，对每个参数进行修正，得到最终参数学习结果；最后，运用经典草地湿润网络模型对提出的方法进行仿真验证。研究结果表明，在小数据集条件下，提出的方法不仅满足了乘性协同约束，而且 KL 散度始终低于其他 2 种方法，但运行时间略高于其他 2 种方法约 1×10^{-3} s，影响甚微。总体上，所提算法的综合性能优于其他 2 种方法。

关键词　贝叶斯网络；参数学习；小数据集；乘性协同约束
DOI　10.3969/j.issn.2097-1915.2023.04.011
中图分类号　TP181　　**文献标志码**　A　　**文章编号**　2097-1915(2023)04-0069-08

BN Parameter Learning Based on Modified Multiplicative Collaborative Constraints with Small Data Sets

REN Zhifang，CHEN Haiyang*，HUAN Xiaomin，SHANG Shanshan
（School of Information and Electronics，Xi'an Polytechnic University，Xi'an 710048）

Abstract　Aimed at the problems that under some specified conditions，obtaining sufficient samples is so difficult that the accuracy of the BN parameters learned by the maximum likelihood estimation algorithm is often low，and multiple parent nodes collaboration to influence constraints is involved in some areas of practical application，a BN parameter learning method based on the modified multiplicative co-constraint under small data sets is proposed by drawing on the idea of PAVA order-preserving regression algorithm. First，the paper is to determine whether the parameters in the multi-parent part of the known sample data meet the needs of the multiplicative collaborative constraint. Secondly，both the left and right sides not to meet the needs of the multiplicative co-constraint are divided into wholes，and adjusted separately by using the PAVA algorithm. And then，for the adjusted whole，three correction methods with different weights are given to correct each parameter according to the amount of sample data corresponding to the combined

收稿日期：2022-12-05
基金项目：国家自然科学基金项目(51905405)
作者简介：任智芳(1996－)，女，陕西宝鸡人，硕士生，研究方向为贝叶斯网络。E-mail:770251684@qq.com
通信作者：陈海洋(1967－)，男，陕西西安人，副教授，博士，研究方向为人工智能、贝叶斯网络。E-mail:chy_00@163.com

引用格式：*任智芳，陈海洋，环晓敏，等. 小数据集下基于修正乘性协同约束的 BN 参数学习[J]. 空军工程大学学报，2023，24(4)：69-76.*
REN Zhifang，CHEN Haiyang，HUAN Xiaomin，et al. BN Parameter Learning Based on Modified Multiplicative Collaborative Constraints with Small Data Sets[J]. Journal of Air Force Engineering University，2023，24(4)：69-76.

state of different parent nodes, and gain mean final parameter learning result. Finally, the proposed method is validated by simulation using a classical grassland wetting network model. The experimental results show that the proposed method not only meets the needs of the multiplicative cooperation constraint under small data set conditions, but also the KL scatter is always lower than the other 2 methods in addition to that the running time is slightly higher than that of the other 2 methods by about 1×10^{-3}s with minimal impact. Generally speaking, the proposed algorithm is superior to the other 2 methods in the comprehensive performance.

Key words Bayesian networks; parameter learning; small dataset; multiplier cooperation

贝叶斯网络(Bayesian networks, BN)是一种将有向无环图结构和条件概率表相结合来表示不确定性知识的网络模型[1-2],已成为解决不确定性问题的有效工具。目前,被广泛应用于医疗分析[3-4]、故障诊断[5-6]、军事智能[7-11]等领域。BN 参数学习是指对结构中各节点条件概率的学习即条件概率表,当有充足的数据时,经典的最大似然估计算法(maximum likelihood estimation,MLE)[12]能够有效地满足参数学习。而对于一些特定情况,收集大量可靠的训练数据是非常困难的,如军事领域中的威胁评估与目标识别、医疗中的罕见疾病、环境风险中的地质灾害等,想要采集到充足的数据都十分不易。若此时,依旧使用 MLE 算法进行参数学习,可能导致估算出的 BN 参数精度较低,从而对后续的推理产生影响,导致最终难以做出较为准确的决策。

针对上述因数据量不充足导致参数学习精度低、推理结果差等问题,基于小样本数据集对 BN 参数学习展开研究显得十分必要。目前,已有许多学者针对小数据集下的 BN 参数学习进行了研究,同时取得了一些研究成果。现有方法大致可分为两类。第一类是基于约束优化模型[13-16]的方法,将参数学习问题当成一种带有约束的模型优化问题,如惩罚函数模型[15]和凸优化模型[16],通过这些模型使得参数满足某种约束。此类方法可以结合几乎所有的约束类型,但缺点是计算量较大。第二类是基于贝叶斯估计[17-20]的方法,将定性专家约束以虚拟样本集的形式转化为定量的参数先验信息,再结合贝叶斯估计算法求解参数。文献[17]针对参数的近似等式约束这种较为常见的专家先验知识,提出采用正态分布对其进行表示来获取超参数,但对约束区间内的参数取值未展开讨论。文献[18]将模糊理论引入参数学习,提出一种基于模糊约束的贝叶斯估计方法,但该方法构建模糊超参的过程较为复杂。文献[19]针对定性最大后验概率算法(qualitative maximum a posterior,QMAP)在参数约束数量较多时无法获得满足约束的参数,设计了一种约束区域中心点的计算方法改进 QMAP 算法,但参数学习精度相较于 QMAP 方法略差。文献[20]通过分析参数性和非参数性的贝叶斯估计,提出采用 β-二项式分布模型构建约束,有效减少了参数估计值的方差,但引入了小量偏差。此类方法的计算较为简单,学习精度也比较准确,但对参数约束条件的限制要求很高,有一定的局限性。此外,还有基于保序回归模型[21]的方法,此类方法可以与定性影响约束相结合对参数进行学习。

通过分析已有数据条件下 BN 参数学习的研究可知,现有文献中基于约束优化模型的方法研究最多的是对单父节点的约束优化问题,对于多父节点的约束问题涉及甚少。但是,目前军事领域中导弹对海基拦截系统的突防评估、防空威胁评估等实际应用问题中均存在多父节点条件下的协同影响约束问题。因此,本文提出了一种修正乘性协同约束的参数学习方法。乘性协同约束[22]本身就是一种用于解决多父节点条件下约束优化问题的方法。本文通过借鉴 Ad Fleeders[23]提出的 PAVA(pool adjacent violators algorithm)思想对乘性协同约束条件下 BN 参数学习进行优化。首先,判断多父节点部分的参数是否满足乘性协同约束,将不满足约束的参数划分成模块,用 PAVA 算法进行调整;然后,根据不同父节点组合状态所对应的样本数据量,给出 3 种权值不同的参数校正方法;最后,利用经典草坪湿润模型对本文方法进行仿真验证。

1 相关理论

1.1 贝叶斯网络中的参数学习

BN 参数学习中,对于每个节点 X_i 和它的父节点集合 $Pa(X_i)$都对应一个条件概率表 $P(X_i|Pa(X_i))$。记网络中条件概率为 θ_{ijk},$\theta_{ijk}=P(X_i=k|Pa(X_i)=j)$。对于具有 n 个节点变量的网络,参数估计可以表示为式(1)所示的对数似然函数最大化问题。

$$l(\theta \mid D)=\sum_{i=1}^{n}\sum_{j=1}^{q_i}\sum_{k=1}^{r_i}N_{ijk}\log\theta_{ijk} \tag{1}$$

式中：N_{ijk} 为样本数据集中满足 $X_i=k$ 且其父节点 $Pa(X_i)=j$ 的样本个数；θ_{ijk} 的最大似然取值满足式(2)时，似然函数值最大。

$$\theta_{ijk}^{*}=\begin{cases}\dfrac{N_{ijk}}{\sum_{k=1}^{r_i}N_{ijk}}, & N_{ijk}>0\\ \dfrac{1}{r_i}, & \text{其他}\end{cases} \tag{2}$$

式中：r_i 是节点 i 的状态总数。

1.2　乘性协同约束的数学模型

协同影响约束是一种非单调定性影响约束，是相对于单调影响约束而提出的[22-23]。如节点 A 对节点 C 的单调影响约束的符号仅取决于节点 A 对 C 的影响，而节点 A 对节点 C 的非单调定性影响约束的符号还取决于其他节点对 C 的影响。若节点 A 是节点 C 的唯一父节点且存在正单调定性影响约束时，则其正定性影响的状态可用其小写形式 a，c 表示如下：

$$P(c\mid a)\geqslant P(c\mid \bar{a}) \tag{3}$$

当节点 C 的父节点除节点 A 之外还存在其它节点时，意味着其父节点的数量恒大于 1。此时，子父节点间的单调定性影响约束不成立，需将其转化为非单调定性影响约束。若节点 A 和 B 都是节点 C 的父节点且存在正定性影响约束时，则其状态可用其小写形式 a，b，c 表示如下：

$$P(c\mid a,b)\geqslant P(c\mid \bar{a},b)\text{ 且 }P(c\mid a,\bar{b})\geqslant P(c\mid \bar{a},\bar{b}) \tag{4}$$

正乘性协同约束[22,24]描述的是父节点个数大于 1 条件下的所有父节点对其子节点的联合影响约束。设某一贝叶斯网络中存在 3 个变量 A，B，C 且均为离散的布尔值。它们之间存在节点 A，B 均为 C 的父节点的关系，则其正乘性协同约束可表示为：节点 A，B 对节点 C 的共同影响大于节点 A，B 分别单独对节点 C 的影响的乘积，具体可表示为：

$$P(c\mid a,b)P(c\mid \bar{a},\bar{b})\geqslant P(c\mid \bar{a},b)P(c\mid a,\bar{b}) \tag{5}$$

同理，负乘性协同可表示为：

$$P(c\mid a,b)P(c\mid \bar{a},\bar{b})\leqslant P(c\mid \bar{a},b)P(c\mid a,\bar{b}) \tag{6}$$

1.3　基于 PAVA 的参数学习算法

保序回归算法的基本原理是将得到的变量与已知变量的序关系进行比较，当得到的变量中有元素不满足已知序关系时，需进行修正，使其最终得到的数据集为一组非递减序列，且使得预测值与真实值的误差最小。通常，在引入权重时，需要使用 PAVA 算法来解决保序回归问题。首先，对每个模块中的样本序列运用 PAVA 算法使得模块内样本有序；然后，对模块间（即整体样本序列）运用 PAVA 算法使得所有样本有序。将 PAVA 算法思想运用到 BN 参数学习中如下：

对于某一节点 X_i，设其状态取值为 k 时，相应的各父节点组合状态取值可表示为 j，而且 $j\in[1,s]$ 时对应的各参数大小关系完全已知。此时，用如下方法对参数 θ_{ijk} 进行计算：

假设 $j_1\leqslant j_2\leqslant j_3\leqslant\cdots\leqslant j_s$ 为父节点取不同状态值时的序关系，式中 j_s 表示 $j=s$。此时，不同状态值所对应的参数大小关系可表示如下：

$$\theta_{ij_1k}\leqslant\theta_{ij_2k}\leqslant\theta_{ij_3k}\leqslant\cdots\leqslant\theta_{ij_sk} \tag{7}$$

当参数之间的大小关系确定之后，把最大似然估计算法学习得到的参数作为初始值，将其与已知的大小关系进行一一比较，若符合，则表示满足约束；若不符合，则需要利用 PAVA 算法对违反部分进行调整。具体计算过程如下：

首先，将网络中的待求参数 $\theta_{ijk}(j=1,2,\cdots,s)$ 划分成模块 $m_1,m_2,\cdots,m_s$，并且每个模块均包含参数值 $v(m_s)$ 和权值 $\omega(m_s)$ 两部分，$j\in[1,s]$，具体表示如下：

$$\omega(m_s)=N_{j_s},v(m_s)=\theta_{ij_sk} \tag{8}$$

其次，检查各模块是否满足式(9)和式(10)：

$$\begin{gathered}Pav[b,d]>Pav[d+1,t]\\ b\in[1,s],d\in[b,s],t\in[d+1,s]\end{gathered} \tag{9}$$

$$Pav[b,d]=\frac{\sum_{s=b}^{d}\omega(m_s)v(m_s)}{\sum_{s=b}^{d}\omega(m_s)} \tag{10}$$

若不满足，则转向下一步；若满足，则说明当前进行比较的参数模块违背了式(7)中的约束，需要更新参数的模块信息。具体操作为：将违背约束的 2 个模块集合并为 1 个模块集并对其加权平均得到新模块集，公式表示如下：

$$\omega[b,t]=\omega[b,d]+\omega[d+1,t] \tag{11}$$

$$Pav[b,t]=\frac{\begin{bmatrix}\omega[b,d]Pav[b,d]+\\ \omega[d+1,t]Pav[d+1,t]\end{bmatrix}}{\omega[b,d]+\omega[d+1,t]} \tag{12}$$

然后，判断所得参数是否均满足式(7)的约束。若不满足，则根据以上步骤进行迭代直至满足约束停止；若满足，则用 $\theta_{ijk}=Pav[b,t]$ 更换新模块集 $[b,t]$ 中的所有参数。最后，当各父节点组合状态取值的序关系完全已知时，所有计算即刻终止。

2 基于乘性协同约束的 BN 参数学习

2.1 算法基本思想

在小数据集条件下，MLE 算法常常无法学习得到精确参数，而专家经验知识往往可以提供较为准确的约束信息。因此，通过加入专家经验或领域知识对参数进行约束优化，来提高参数的学习精度。本文中，首先，利用最大似然估计算法学习得到网络参数并令其为初始参数用于后续学习，同时假设从专家经验知识中得到的约束均正确。然后，将初始参数划分为单个的模块并判断其是否满足乘性协同约束；若不满足，则需要利用 PAVA 算法对相应模块进行调整；若满足，则将其作为最后的学习结果。最后，将调整之后的参数通过不同的平均策略进行再次校正，这是该算法的核心所在。该算法的特点是：仅对违反了乘性协同约束的参数进行调整并进一步校正，对于未违反约束的参数就将其作为最后学习结果；在充分利用小数据集有限样本量的同时又对违反约束的参数加以校正，两方面相结合使得参数学习的结果更加准确。

2.2 3 种参数校正方法

通过借鉴 PAVA 算法思想，首先对不满足约束的参数进行调整，然后利用不同的平均策略结合乘性协同约束给出了 3 种不同的方法对参数进行校正。3 种校正方法的异同主要在于求取 mean 值和进一步修正参数时的权值分配。其中：方法 1 和方法 2 先选用不同的计算方法求 mean 值，再进一步采用相同的校正公式修正参数；方法 2 和方法 3 先采用相同的计算方法求 mean 值，再进一步选用不同的校正公式修正参数。在小数据集条件下，若网络中的参数变量 A,B,C 符合 1.2 节提到的正乘性协同约束，则具体计算方法如下，负乘性协同约束同理可得。

方法 1：

1）求取网络参数后，根据参数样本量赋予每个参数相应的权值，将违反乘性协同约束的部分，借鉴 PAVA 思想加权平均得到 mean 值：

$$\text{mean}=\frac{\begin{bmatrix}(N_1+N_4)P(c|a,b)P(c|\bar{a},\bar{b})+\\(N_2+N_3)P(c|\bar{a},b)P(c|a,\bar{b})\end{bmatrix}}{N_1+N_2+N_3+N_4}$$

2）引入比例系数 f_1、f_2，按相应的权值分别对每个参数进行校正，样本量大的则权值大一些，样本量小的则权值小一些，具体方法如下：

$$f_1=\text{mean}/[P(c|a,b)P(c|\bar{a},\bar{b})]$$

$$f_2=[P(c|\bar{a},b)P(c|a,\bar{b})]/\text{mean}$$

$$P^N(c|a,b)=P(c|a,b)f_1^{\frac{N_1}{N_1+N_4}}$$

$$P^N(c|\bar{a},\bar{b})=P(c|\bar{a},\bar{b})f_1^{\frac{N_4}{N_1+N_4}}$$

$$P^N(c|\bar{a},b)=P(c|\bar{a},b)f_2^{-\frac{N_2}{N_2+N_3}}$$

$$P^N(c|a,\bar{b})=P(c|a,\bar{b})f_2^{-\frac{N_3}{N_2+N_3}}$$

方法 2：

1）求取网络参数后，赋予每个参数相同的权值，将违反乘性协同约束的部分，借鉴 PAVA 思想加权平均得到 mean 值

$$\text{mean}=\frac{\begin{bmatrix}P(c|a,b)P(c|\bar{a},\bar{b})+\\P(c|\bar{a},b)P(c|a,\bar{b})\end{bmatrix}}{2}$$

2）根据样本量赋予每个参数相应的权值，引入比例系数 f_1、f_2，按不同的权值分别对每个参数进行校正，具体方法如下：

$$f_1=\text{mean}/[P(c|a,b)P(c|\bar{a},\bar{b})]$$

$$f_2=[P(c|\bar{a},b)P(c|a,\bar{b})]/\text{mean}$$

$$P^N(c|a,b)=P(c|a,b)f_1^{\frac{N_1}{N_1+N_4}}$$

$$P^N(c|\bar{a},\bar{b})=P(c|\bar{a},\bar{b})f_1^{\frac{N_4}{N_1+N_4}}$$

$$P^N(c|\bar{a},b)=P(c|\bar{a},b)f_2^{-\frac{N_2}{N_2+N_3}}$$

$$P^N(c|a,\bar{b})=P(c|a,\bar{b})f_2^{-\frac{N_3}{N_2+N_3}}$$

方法 3：

1）求取网络参数后，根据参数样本量赋予每个参数相同的权值，将违反乘性协同约束的部分，借鉴 PAVA 思想加权平均得到 mean 值

$$\text{mean}=\frac{\begin{bmatrix}P(c|a,b)P(c|\bar{a},\bar{b})+\\P(c|\bar{a},b)P(c|a,\bar{b})\end{bmatrix}}{2}$$

2）引入比例系数 f_1、f_2，按相同的权值对每个参数进行校正，具体方法如下：

$$f_1=\text{mean}/[P(c|a,b)P(c|\bar{a},\bar{b})]$$

$$f_2=[P(c|\bar{a},b)P(c|a,\bar{b})]/\text{mean}$$

$$P^N(c|a,b)=P(c|a,b)f_1^{\frac{1}{2}}$$

$$P^N(c|\bar{a},\bar{b})=P(c|\bar{a},\bar{b})f_1^{\frac{1}{2}}$$

$$P^N(c|\bar{a},b)=P(c|\bar{a},b)f_2^{-\frac{1}{2}}$$

$$P^N(c|a,\bar{b})=P(c|a,\bar{b})f_2^{-\frac{1}{2}}$$

式中：N_1,N_2,N_3,N_4 分别表示父节点取值状态为 (a,b)、$(\bar{a},b)$、$(a,\bar{b})$、$(\bar{a},\bar{b})$时的样本数据量，一般情况下，3 种参数校正方法均可使用；若不同父节点组合状态的统计量差异较大时可优先采用方法 1；若 $N_1\neq N_4$，$N_2\neq N_3$ 时可优先采用方法 2；若 $N_1\approx N_4$，$N_2\approx N_3$ 时可优先采用方法 3；P^N 表示校正后的参数。

假设贝叶斯网络中共有 D 个节点，其中隐藏节点为 n 个，节点的最大状态数为 N，在使所有参数都满足式(7)约束的迭代过程中，每种参数校正方法所需运算次数分别为 a_1,a_2,a_3，从而计算出本文方法 1～方法 3 的复杂度分别为：$O(a_1(2D-n)N^D)$、$O(a_2(2D-n)N^D)$、$O(a_3(2D-n)N^D)$。

2.3　算法描述

在小数据集条件下，若网络中的参数变量符合 1.2 节给出的乘性协同约束数学模型时，将结合 PAVA 保序回归算法调整并利用不同校正方法对贝叶斯网络参数进行学习。

算法流程如图 1 所示。

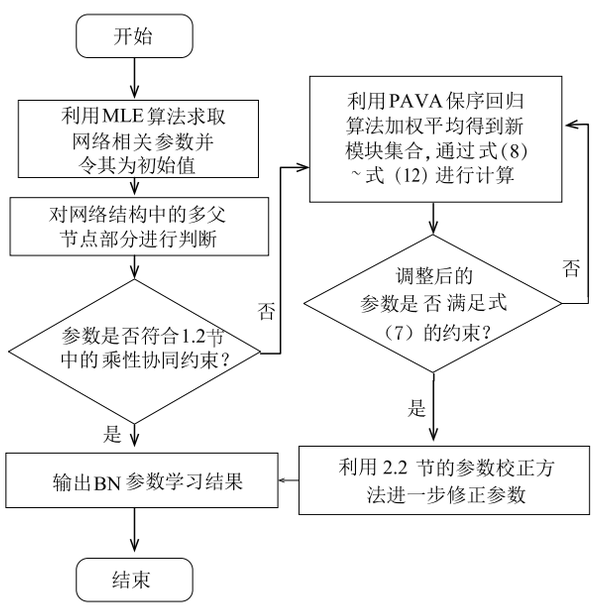

图 1　算法流程图

算法具体步骤如下：

步骤 1　对已知小样本数据利用最大似然估计算法进行学习，将其得到的相关参数作为后续学习中的网络初始参数。

步骤 2　判断已知样本数据中多父节点部分的参数是否满足乘性协同约束，若满足，则转向步骤 6，若不满足，则返回步骤 3。

步骤 3　对于不满足乘性协同约束的参数，利用 PAVA 保序回归算法对其相邻模块进行合并之后加权平均得到新模块集，通过式(8)～式(12)进行计算。

步骤 4　判断调整之后的参数是否满足式(7)的约束，若满足，则转向步骤 5，若不满足，则依据步骤 3 进行迭代直至满足停止。

步骤 5　针对调整后的整体(参数之积)，采用 2.2 节的校正方法对每个参数进行相应的修正。

步骤 6　所得参数即为最终的参数学习结果。

3　仿真结果及分析

3.1　实验条件

本次仿真采用 Windows 10 系统以及 Matlab R2014a 作为研究平台。为了验证分析本文提出的算法性能，选用如图 2 所示的经典草坪湿润网络模型，该网络模型由 4 个节点和 4 条边构成，其中节点 C 表示天气是否多云，节点 S 表示是否需要洒水车进行洒水，节点 R 表示天气是否下雨，节点 W 表示草坪是否湿润；节点间的有向边由父节点指向其子节点，表示节点间的相互依赖关系。网络模型的选取虽然比较简单，但是模型中不仅包含有多父节点关系，而且涉及 BN 结构学习中最基础的顺连、汇连和分连 3 类结构。因此，草坪湿润网络模型可以作为一种典型网络用以评价 BN 参数学习方法的优劣性。

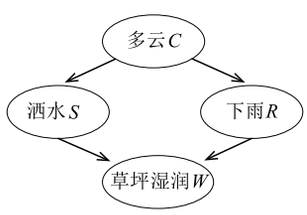

图 2　草坪湿润模型

严格来说，数据的“充足”或“不充足”是一个相对量，大多领域都面临数据不充足的问题，但是并未有领域明确对小数据集给出定义。在贝叶斯网络学习中，衡量一个网络所需数据量的多少，需要考虑多方面因素，如网络结构复杂度、学习精度、最大父节点数目等。

当 BN 的网络结构已知且所有节点均是离散值的条件下，可以通过最大似然估计计算 BN 网络所需样本数据量[25]。计算公式如下：

$$M\geqslant\frac{288\times 2^K}{\varepsilon^2}\ln^2\left(1+\frac{3}{\varepsilon}\right)\ln\left(\frac{18n\times 2^K\ln\left(1+\frac{3}{\varepsilon}\right)}{\varepsilon\delta}\right)\tag{13}$$

式中：M 表示样本数量；K 表示节点的最大状态数；ε 表示每个节点对应的参数的 KL 误差；n 表示网络

节点数量；δ 为置信度。

取 $\varepsilon=5$，$K=2$，$\delta=0.05$ 时，当样本数据量小于通过式(13)计算所得的样本数据量的数据集称为小数据集。

本次研究中，首先采用 MLE 算法、改进的 QMAP 方法[26]、本文方法 1、方法 2、方法 3 依次取不同的样本量对 BN 参数进行学习，然后分别以 KL 散度和运行时间作为指标，分析和对比算法的准确度和复杂度。仿真网络的部分真实参数如表 1 所示。

表 1 仿真网络的部分真实参数

	$S=0$，$R=0$	$S=0$，$R=1$	$S=1$，$R=0$	$S=1$，$R=1$
$W=1$	0.4	0.6	0.6	0.95
$W=0$	0.6	0.4	0.4	0.05

3.2 仿真分析

假设图 2 的模型中 R，S，W 之间的参数满足正乘性协同约束，其中变量取 1 为事件发生，取 0 为不发生。具体表示为：

$$P(W=1|R=1,S=1)\times P(W=1|R=0,S=0)\geqslant P(W=1|R=1,S=0)\times P(W=1|R=0,S=1)$$

取样本数据量为 35 时，MLE 方法和本文方法得到的仿真结果如表 2～表 5 所示。

表 2 MLE 学习得到参数

	$S=0$，$R=0$	$S=0$，$R=1$	$S=1$，$R=0$	$S=1$，$R=1$
$W=1$	0.777 8	0.6	0.5	1
$W=0$	0.222 2	0.4	0.5	0

表 3 方法 1 学习得到的参数

	$S=0$，$R=0$	$S=0$，$R=1$	$S=1$，$R=0$	$S=1$，$R=1$
$W=1$	0.584 5	0.653 1	0.660 7	0.738 2
$W=0$	0.415 6	0.346 9	0.339 3	0.261 8

表 4 方法 2 学习得到的参数

	$S=0$，$R=0$	$S=0$，$R=1$	$S=1$，$R=0$	$S=1$，$R=1$
$W=1$	0.582 0	0.650 3	0.657 4	0.734 6
$W=0$	0.418 0	0.349 7	0.342 6	0.265 4

表 5 方法 3 学习得到的参数

	$S=0$，$R=0$	$S=0$，$R=1$	$S=1$，$R=0$	$S=1$，$R=1$
$W=1$	0.579 6	0.653 8	0.653 8	0.737 6
$W=0$	0.420 4	0.346 2	0.346 2	0.262 4

从表 2～表 5 大致可知：学习所得参数与真实参数最为接近的是本文方法 3。为进一步比较算法的学习精度，采用 KL 散度来描述学习得到的参数与真实参数之间的拟合程度，所得值越小说明其学习精度越高，KL 散度的计算公式如式(14)所示。在不同数据样本量条件下分别进行 50 次实验并求取平均值，将本文提到的 3 种方法与 MLE 算法、改进的 QMAP 方法进行分析对比。所得 KL 散度结果图如图 3、图 4 所示，其中图 3 为几种方法的全局显示图，为了更清晰地对比观察 MLE 算法、改进的 QMAP 方法与本文方法的 KL 散度，特将纵坐标 KL 散度数值范围缩小至 0.13～0.35，得到如图 4 所示的局部显示图。

$$KL(P,Q)=\sum_{X}P(X)\log\frac{P(X)}{Q(X)} \tag{14}$$

通过观察分析图 3、图 4 可知：在小样本量条件下，本文方法的 KL 散度均明显小于 MLE 算法和改进的 QMAP 方法的 KL 散度，且本文方法 3 的 KL 散度折线图一直处于最低，这就说明在小数据集条件下，在学习精度方面本文的 3 种方法均优于 MLE 算法和改进的 QMAP 方法。其原因是本文方法基于 MLE 算法之上，通过引入乘性协同约束对样本数据进行调整和修正，既保证小数据集得到充分利用，又克服了 MLE 算法在小数据集条件下学习结果偏差大的不足，同时也有效提高了参数学习的精度。对比本文的 3 种方法可知：随着数据量的不断增加，方法 3 的学习精度一直保持着最高；在数据量较小时，方法 2 的学习精度次之，方法 1 的学习精度最低；在数据量较大时，方法 1 的学习精度次之，方法 2 的学习精度最低。

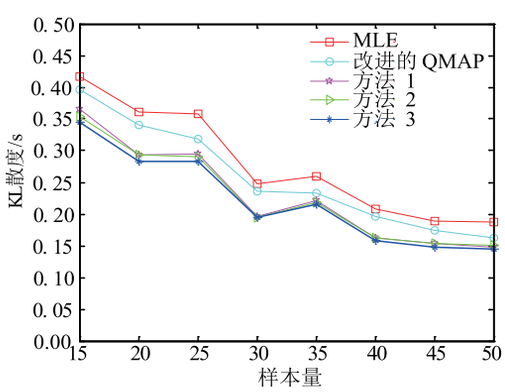

图 3 几种算法的 KL 散度比较

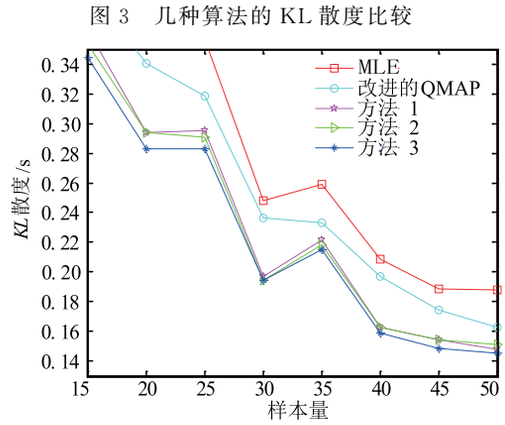

图 4 几种算法的 KL 散度比较(局部放大)

3 种算法的学习均需要真实样本数据作为基础，不同之处在于，本文方法对不满足约束的部分，

先通过 PAVA 算法给予调整，再结合乘性协同约束进行修正，使少量数据充分发挥作用；随着样本数据量的不断扩充，学习到的结果也会更加准确。而改进的 QMAP 方法中加入了虚拟样本，所以在小数据集下，真实样本对算法的学习精度影响较小，从而导致随着真实样本的增加，算法的学习精度提高也较小，整体上呈现下降趋势。但是相较于改进的 QMAP 方法，本文 3 种方法学习到的参数的 KL 精度一直处于最小。因此，本文方法更适用于小数据集条件下的 BN 参数学习，在小数据集条件下取得的学习结果也更加准确。

以上侧重分析了本文提出算法的学习精度，单从学习精度来看，本文方法在样本数据量较小的条件下更占优势。为了进一步验证本文方法的性能，接下来将对算法复杂度进行简要分析。通常，将算法的运行时间作为衡量其复杂度的指标，对此，基于不同样本量对几种算法的时效性进行测试。在测试过程中，将每一组样本数据量各运行 50 次并求取其平均值，得到的运行时间结果图如图 5 所示。

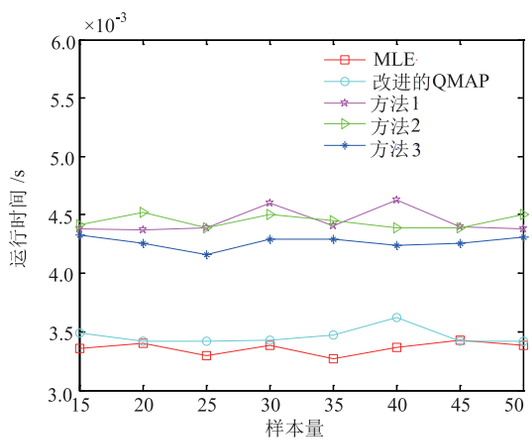

图 5　几种算法的运行时间比较

从图 5 可知：几种算法的时间消耗均较小，其中 MLE 算法和改进的 QMAP 方法的运行速度较快，本文提出的方法 3 次之，方法 1 和方法 2 的运行时间均衡来看相当。主要是由于本文算法除了包含 MLE 算法的计算量之外，还包含调整及修正参数的过程，需要判断出已知样本数据中不满足乘性协同约束的参数，对这部分不满足约束的参数运用保序回归方法进行顺序调整，然后再通过不同的平均策略对每个参数进一步修正。这些步骤的运行均需要消耗时间，这就必然导致算法运行时间的增加。一般情况下，算法的学习精度提升必然会导致运行时间上的增加。本文方法在有效提高参数学习精度的同时也确保了学习结果满足所有约束，因而运行时间上的略微增加可以忽略不计。因此，通过对 KL 散度和运行时间两方面测试结果的综合分析，可以得出本文算法的综合性能好。

4　结语

在小数据集条件下，目前已有的 BN 参数学习方法得到的结果，不仅经常会违反专家约束，并且基于多父节点的参数学习研究比较少。因而本文针对多父节点的参数进行了学习，首先通过借鉴 PAVA 算法的思想对于违反约束关系的样本进行调整使其最终满足非递减的序列关系，然后结合乘性协同约束方法通过不同的平均策略进行具体的计算及修正，使其满足专家约束的同时让有限的小数据量得到充分利用。该方法的优点在于它保证了样本数据在违反乘性协同约束的情况下，先调整违反约束的样本顺序，再运用不同的方法进行参数校正，使其满足专家约束，以达到更准确的学习精度。本文分别从算法的学习精度和运行时间两方面对几种方法进行对比分析。依据仿真结果可以得出：本文提出的 3 种方法在学习精度方面，相比 MLE 算法和改进的 QMAP 方法 KL 散度值较小，并且本文提出的算法中方法 3 的学习精度最高；但在时效性方面，相比 MLE 算法和改进的 QMAP 方法，本文 3 种方法的耗时均略微延长。

综合来看，本文方法为小数据集条件下多父节点贝叶斯网络结构的参数学习提供了一种新思路。在以后的研究中，可以结合其他理论方法来研究小数据集条件下乘性协同约束的融合处理方法。

参考文献

[1] SCANAGATTA M, CORANI G, ZAFFALON M, et al. Efficient Learning of Bounded-Treewidth Bayesian Networks from Complete and Incomplete Data Sets[J]. International Journal of Approximate Reasoning, 2018, 95(4): 152-166.

[2] 杜文静，刘海. 贝叶斯人工智能的概率证成[J]. 科学技术哲学研究，2022，39(4)：16-20.

[3] 徐夏楠，张洪. 基于信息增益的加权贝叶斯插补法及其在心脏病类医疗缺失数据分析中的应用[J]. 复旦学报(自然科学版)，2022，61(3)：335-341,352.

[4] ZHANG T, ZHANG T, LI C, et al. Complementary and Alternative Therapies for Precancerous Lesions of Gastric Cancer: A Protocol for a Bayesian Network Meta Analysis[J]. Medicine, 2021, 100(2):e24249.

[5] 李登峰，林萍萍. 基于 D-S 证据融合和直觉模糊贝叶斯网络双向推理的景区游客拥挤踩踏故障诊断分析[J]. 系统工程理论与实践，2022，42(7)：1979-1992.

[6] AMIN M T, KHAN F, IMTIAZ S. Fault Detection and Pathway Analysis Using a Dynamic Bayesian Network[J]. Chemical Engineering Science, 2019, 195: 777-790.

[7] 夏命辉，王小平，林秦颖，等. 复杂环境下基于动态贝叶斯网络的目标识别[J]. 空军工程大学学报(自然科学版)，2016，17(4)：24-28.

[8] 高天祥，王刚，岳韶华，等. 基于贝叶斯决策理论的NSHV分段建模威胁评估[J]. 空军工程大学学报(自然科学版)，2019，20(1)：60-66.

[9] CHEN L, MA Y P. Simulation Research on Anti-submarine Target Identification and Threat Assessment of Aircraft Carrier Formation[J]. Fire Control & Command Control, 2019, 44(3): 153-158,164.

[10] 高晓光，杨宇. 基于贝叶斯网的舰艇防空威胁评估[J]. 战术导弹技术，2020(4)：47-57,70.

[11] 严惊涛，刘树光，杜梓冰. 贝叶斯网络的对地攻击无人机自主能力评估[J]. 空军工程大学学报，2022，23(4)：92-98.

[12] BONAITI L, GORLA C. Estimation of Gear SN Curve for Tooth Root Bending Fatigue by Means of Maximum Likelihood Method and Statistic of Extremes[J]. International Journal of Fatigue, 2021, 153(12): 106451.

[13] GUO Z G, GAO X G, REN H, et al. Learning Bayesian Network Parameters from Small Data Sets: a Further Constrained Qualitatively Maximum a Posteriori Method[J]. International Journal of Approximate Reasoning, 2017, 91(12): 22-35.

[14] 魏曙寰，曾强，陈砚桥. 基于AHP/D-S证据理论的贝叶斯网络参数学习方法[J]. 海军工程大学学报，2021，33(6)：19-24.

[15] MAMMARELLA M,ALAMO T,LUCIA S,et al. A Probabilistic Validation Approach for Penalty Function Design in Stochastic Model Predictive Control [J]. IFAC Papers on Line, 2020, 53(2): 11271-11276.

[16] AGRAWAL A, BARRATT S, BOYD S. Learning Convex Optimization Models[J]. CAA Journal of Automatica Sinica, 2021, 8(8): 1355-1364.

[17] 柴慧敏，赵昀瑶，方敏. 利用先验正态分布的贝叶斯网络参数学习[J]. 系统工程与电子技术，2018，40(10)：2370-2375.

[18] 茹鑫鑫，高晓光，王阳阳. 基于模糊约束的贝叶斯网络参数学习[J]. 系统工程与电子技术，2023，45(2)：444-452.

[19] 邸若海，李叶，万开方，等.基于改进QMAP的贝叶斯网络参数学习算法[J]. 西北工业大学学报，2021，39(6)：1356-1367.

[20] GRIBOK A, AGARWAL V, YADAV V. Performance of Empirical Bayes Estimation Techniques used in Probabilistic Risk Assessment[J]. Reliability Engineering and System Safety, 2020, 201:106805.

[21] DI R, GAO X, GUO Z. Learning Bayesian Network Parameters under New Monotonie Constraints[J]. Journal of Systems Engineering and Electronics, 2017, 28(6): 1248-1255.

[22] WELLMAN M P. Fundamental Concepts of Qualitative Probabilistic Networks[J]. Artificial Intelligence, 1990, 44(3): 257-303.

[23] FEELDERS A, GAAG L. Learning Bayesian Networks Parameters under Order Constraints[J]. International Journal of Approximate Reasoning, 2006, 42(1-2): 37-53.

[24] RENOOIJ S. Qualitative Approaches to Quantifying Probabilistic Networks [J]. Utrecht University, 2001: 14-21.

[25] KOLLER D, FRIEDMAN N. Probabilistic Graphical Models[J]. International Journal of Intelligent Systems, 2003, 18(2):149-151.

[26] 陈海洋，张静，王露楠，等. 小数据集下基于改进QMAP算法的BN参数学习[J]. 西安工程大学学报，2023，37(1)：126-133.

(编辑：杜娟)