

因果发现算法研究综述

祁慧英¹, 苏恒², 王浩莲³, 苗志轩^{2*}

¹伊犁师范大学电子工程学院, 新疆 伊宁

²伊犁师范大学网络安全与信息技术学院, 新疆 伊宁

³伊犁河谷智能计算研究与应用重点实验室, 新疆 伊宁

收稿日期: 2026年2月28日; 录用日期: 2026年3月27日; 发布日期: 2026年4月7日

摘要

因果发现旨在从观测数据或实验数据中自动化地辨识变量之间的因果关联结构, 是因果推理研究领域的核心议题之一。本文系统性地梳理了因果发现的理论根基、算法分类体系及其演进历程。首先, 阐明因果发现问题的基本形式化框架, 涵盖结构因果模型、有向无环图、因果马尔可夫条件与忠实性假设等核心理论要素; 其次, 依据方法论范式将现有算法划分为基于约束的方法、基于评分的方法、基于函数因果模型的方法、混合策略方法以及基于深度学习的新兴方法五大类别, 逐一剖析其理论动机、代表性算法、技术优势与固有局限; 最后, 对该领域面临的开放性挑战与未来发展趋向进行展望。

关键词

因果发现, 结构因果模型, 有向无环图, 基于约束的方法, 基于评分的方法, 基于混合的方法

A Survey on Causal Discovery Algorithms

Huiying Qi¹, Heng Su², Haolian Wang³, Zhixuan Miao^{2*}

¹School of Electronic Engineering, Yili Normal University, Yining Xinjiang

²School of Cybersecurity and Information Technology, Yili Normal University, Yining Xinjiang

³Key Laboratory of Intelligent Computing Research and Application in Yili River Valley, Yining Xinjiang

Received: February 28, 2026; accepted: March 27, 2026; published: April 7, 2026

Abstract

Causal discovery aims to automatically identify the underlying causal relationships among variables from observational or experimental data, and has become a central topic in the field of causal inference. This paper provides a systematic review of the theoretical foundations, algorithmic taxonomy, and evolutionary development of causal discovery methods. First, the fundamental formal framework

*通讯作者。

of the causal discovery problem is introduced, including core theoretical components such as structural causal models, directed acyclic graphs, the causal Markov condition, and the faithfulness assumption. Second, according to methodological paradigms, existing algorithms are categorized into five major classes: constraint-based methods, score-based methods, functional causal model-based methods, hybrid approaches, and emerging deep learning-based methods. For each category, the theoretical motivation, representative algorithms, technical advantages, and inherent limitations are systematically analyzed. Finally, the paper discusses the open challenges faced by the field and outlines potential directions for future research.

Keywords

Causal Discovery, Structural Causal Models, Directed Acyclic Graphs, Constraint-Based Methods, Score-Based Methods, Hybrid Methods

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

本模块探究事物之间的因果联系是科学研究的基本诉求。无论在自然科学、社会科学还是工程技术领域,诸多关键问题——如基因表达调控网络的刻画、经济政策效应的识别与评估、以及疾病发生机制的阐明——本质上都可归结为对变量之间因果结构的辨识、建模与解释[1]。随机对照实验在确证因果关联方面长期占据范式地位,被公认为该领域的“金标准”。然而,由于伦理约束、高昂成本及技术实现路径的局限,主动干预实验在诸多现实场景中往往难以开展。在此背景下,如何在无干预条件下仅凭借被动观测数据解析深层因果逻辑,即“因果发现”(Causal Discovery)命题,已跃升为统计学、计算机科学与人工智能深度融合的研究制高点[2]。

因果发现的学术渊源可追溯至 20 世纪初期统计学家对相关性与因果性之间本质差异的哲学辨析。Wright 于 1921 年提出的路径分析方法开创了以图模型表征因果关系的先河[3]。然而,因果发现作为一个系统性的算法研究方向,其真正的理论奠基期当数 20 世纪 80 年代至 90 年代。在此阶段, Pearl、Spirtes 及 Glymour 等学者将有向无环图(DAGs)引入因果表征领域,系统地构建了因果图模型这一理论大厦。随着 PC 与 FCI 等具备奠基意义的算法问世,因果发现的研究范式实现了从经验驱动向形式化逻辑的深刻转型。这一系列开创性成果不仅为因果推断赋予了严密的计算属性,更促成了其从古典哲学思辨向现代实证科学体系的逻辑跃迁[4]。

因果发现技术的爆发式增长,使得现有的综述成果难以全面捕捉其最新的动态脉络,因此,构建一套涵盖从理论奠基到现代演进的全景式知识图谱显得尤为紧迫。本研究的边际贡献体现在:一是对因果发现的理论渊源进行溯源式清理,以消除跨学科背景下的概念歧义,形成统一的分析语境;二是立足于方法论范式的差异性,重构算法谱系分类,并深入剖析代表性算法在复杂场景下的鲁棒性与适用特征;三是在审视现有研究局限的基础上,凝练亟待攻克的关键科学问题,为后续的理论突破与工程实践指引航向。

2. 理论基础与核心定义

2.1. 贝叶斯网络

贝叶斯网络作为刻画多变量联合概率分布的概率图模型,为因果关系的表征与推断提供了严谨的数

学框架[5]。形式化地,一个贝叶斯网络定义为二元组 $B = \langle G, P \rangle$, 其中 $G = (V, E)$ 为有向无环图, 节点集合 $V = \{X_1, X_2, \dots, X_n\}$ 对应观测变量, 有向边集合 E 编码变量间的直接依赖关系; P 表示定义于 V 上的联合概率分布[5]。

贝叶斯网络通过局部马尔可夫性质确立了变量间的依赖拓扑, 规定了每个节点在父节点条件下的局部独立性(即排除非后代节点的关联影响)。这种结构化的语义约束为联合分布的高维降维提供了数理依据, 从而允许我们将全局联合概率分解为一系列局部条件分布的乘积叠加:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pa(X_i))$$

其中 $Pa(X_i)$ 表示节点 X_i 在图 G 中的父节点集合。这种因子化解将高维联合分布的表示复杂度从指数级降至多项式级, 为基于观测数据的结构学习奠定了计算基础[5] [6]。

2.2. 条件独立性与 d-分离

设 U 为随机变量集合, X 、 Y 和 Z 为 U 中三个两两不交的变量子集。若对于任意满足 $P(Z) > 0$ 的取值, 均有 $P(X \perp Y | Z) \Leftrightarrow P(X, Y | Z) = P(X | Z)P(Y | Z)$, 则称 X 与 Y 在给定 Z 条件下独立, 记为 $X \perp Y | Z$ 。

在有向无环图中, 条件独立性可通过 d-分离准则进行图结构层面的判定。该准则基于三种基本连接模式定义信息传递的阻断条件: 在链式结构 $X \rightarrow Z \rightarrow Y$ 与分叉结构 $X \leftarrow Z \rightarrow Y$ 中, 中间节点 Z 被观测时阻断 X 与 Y 之间的依赖传递; 在对撞结构 $X \rightarrow Z \leftarrow Y$ 中, 中间节点 Z 未被观测时 X 与 Y 相互独立, 而 Z 或其后代被观测时将激活二者之间的依赖关系[5]。

d-分离准则构成了从图拓扑逻辑演算概率独立性的形式化语义。在忠实性假设的约束下, 概率分布中蕴含的所有条件独立性均可被图结构的 d-分离属性所穷尽, 从而确立了图模型与统计特征间的同构映射。这种结构与分布的深度一致性, 为约束类算法在实证数据中反演因果架构奠定了公理化基石[7]。

2.3. 马尔可夫毯

马尔可夫毯主要由三部分组成, 明确界定各组成部分的含义是后续 MB 学习的基础: 一是父节点 (Parents), 指在因果图中直接指向目标变量 T 的变量; 二是子节点 (Children), 指在因果图中被目标变量 T 直接指向的变量; 三是配偶节点 (Spouses), 指子节点的父节点中除目标变量 T 以外的其他变量。其中, 父节点与子节点共同构成 PC 集合 (Parents + Children), 是 MB 学习中优先识别的核心部分[4]。

定义 1 (马尔可夫毯, MB) 假定贝叶斯网络满足忠实性条件, 则网络中任意节点的马尔可夫毯是独一无二的, 它包括该节点的直接前驱(父亲节点)、直接后继(子节点)以及配偶节点(即与该节点共同作为父节点的其他节点) [8]。

基于忠实性假设, 贝叶斯网络中任意节点 T (其中 $T \in U$) 的马尔可夫毯构成可形式化界定如下(参见图 1): 该集合涵盖该节点的直接前驱集合 $P(T) = \{A, B\}$ 、直接后继集合 $C(T) = \{C, D\}$, 以及经由共同子节点与之形成依赖关系的配偶节点集合 $SP(T) = \{E, F\}$ 。据此, 节点 T 的马尔可夫毯可严格表述为上述三类邻域节点的并集, 即:

$$MB(T) = P(T) \cup C(T) \cup SP(T) = \{A, B, C, D, E, F\}$$

该集合构成了刻画节点 T 局部概率特性的最小变量集合, 在给定马尔可夫毯条件下, T 与网络中其余所有变量条件独立。

这一定义表明, 马尔可夫毯具有双重特性: 充分性——给定马尔可夫毯后, 目标变量与其余所有变量条件独立; 最小性——马尔可夫毯的任何真子集都不具备完全的信息屏蔽能力。

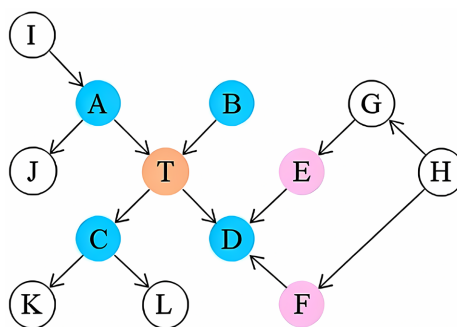


Figure 1. An illustration of a compact Bayesian network
图 1. 小型贝叶斯网络示例

定义 2 (忠实性): 对于一个贝叶斯网络 $\text{BN}\langle U, G, P \rangle$, U 表示网络中的变量全集, G 表示定义在 U 上的有向无环图, P 表示联合概率分布。若对任意 $X, Y \in U$ 以及任意 $S \subseteq U \setminus \{X, Y\}$, 均满足 $X \perp_p Y | S \Leftrightarrow X$ 与 Y 在 G 中被 S -分离, 则称联合概率分布 P 对图 G 满足忠实性[4]。

若要实现对节点间依赖关系的深度解构, 忠实性假设的引入必不可少, 它确保了贝叶斯网络的连接模式能够精确镜像变量间的联合概率分布。在这种严丝合缝的对应关系下, 复杂的条件独立性得以通过图形化手段进行精准刻画, 为关联机制的进一步探究扫清了逻辑障碍。同时, 这一假设在界定目标变量的马尔可夫毯时具有决定性意义, 保障了其在数学推导上的唯一存在。

定义 3 (V-结构) 在一个 DAG 中, 如果有来自 Q 和 W 的两个边分别指向节点 X , 但 Q 和 W 为非相邻节点, 即 $(Q \rightarrow X \leftarrow W)$ (箭头表示节点之间边和边的方向), 则节点 Q 、 X 和 W 的三元结构形成 V-结构[8]。

探究 V-结构的拓扑演化对马尔可夫毯的识别具有深远的理论意义。一方面, V-结构作为锁定“配偶节点”的核心判据, 补全了马尔可夫毯中除父节点与子节点之外的关键成员, 构筑了变量间局部独立性的完备逻辑闭环。另一方面, V-结构是贝叶斯网络结构学习中实现边定向的先验约束, 它通过碰撞诱导的依赖关系变化, 为无向路径的极量化提供了拓扑证明。正是这种双重属性, 确立了 V-结构在因果推断算法体系中的战略地位。

3. 因果发现算法

本节围绕三类主流范式对因果发现算法进行系统阐释: 基于约束(constraint-based)、基于评分(score-based)与基于混合(hybrid)。三者分别从“条件独立关系约束”“全局模型拟合优度”与“两者互补融合”的角度出发, 在统计假设、可识别性边界与计算策略上各具优势与局限。为便于比较, 本文在每一小节均从核心思想、典型算法与关键技术点、适用场景与局限三个层面展开。为保证论述的结构性与可比性, 后续各小节将分别从基本原理、典型算法与技术特征、适用场景及其局限等方面展开系统讨论。

3.1. 基于约束的因果发现

基于约束的因果发现范式将统计层面的条件独立性(CI)视为反演因果拓扑的核心原语, 通过解析观测样本中的独立性逻辑来逆向溯源潜在的因果图谱。其理论构建于马尔可夫性与忠实性双重公理之上: 前者确立了由图结构向独立性分布的逻辑蕴含路径, 后者则保障了统计独立性对图论分离属性的镜像还原。在此框架下, 算法将 CI 模式转化为强化的结构约束, 并经由启发式推理规则实现变量关联及其极性的系统化恢复。在满足因果充分性(Causal Sufficiency)的前提下, 输出结果通常收敛于代表马尔可夫等价类的 CPDAG; 而针对含有潜变量或选择偏倚的复杂场域, 则需诉诸 PAG 来刻画具备可识别性的祖先关

系及拓扑不确定性[4] [9] [10]。

在约束驱动算法谱系中，PC 算法凭借其在无潜变量场景下的基准地位而备受关注[11]。该算法通常采用“两阶段”结构学习范式：初阶段通过递增条件的阶数进行独立性剪枝，旨在从完全图中剔除虚假关联以萃取因果骨架；进阶阶段则聚焦于碰撞体(v-structure)的逻辑识别与定向传播，最终将无向骨架映射为表征马尔可夫等价类的 CPDAG。当研究场域延伸至包含混杂因子或选择偏倚的复杂环境时，以 FCI 及其衍生算子为代表的广义框架展现出更强的鲁棒性[4] [13]。此类方法通过引入精细化的分离集检索机制与强化定向准则，在隐变量干扰下最大化地还原具备可识别性的祖先关联，并利用 PAG 形式精确界定结构的不确定性边界。针对高维环境下条件集膨胀诱发的“维度灾难”，学术界近年来的优化路径聚焦于阶数受限搜索、局部邻域启发式算子、并行化架构及稳定性选择(Stability Selection)等策略，旨在显著增强算法的大规模任务适应性。表 1 对此类具有代表性的约束方法进行了体系化归辑与性能评析。

Table 1. Constraint-based causal representative discovery algorithms: classification, core mechanisms, advantages and disadvantages

表 1. 基于约束的因果代表发现算法分类、核心机制及优劣势对比

代表算法	核心机制	优势	局限性
PC [11]/IC [14]	逐步增加 CI 检验；骨架学习 + V 结构定向。	理论完备，多项式复杂度(稀疏图)。	依赖因果充分性与无环假设；对变量顺序敏感。
IAMB 系列[15]-[17]/GSMB [18]	增长-收缩策略	适合局部 MB 发现。	难以区分父子/配偶；假阳性难以彻底剔除。
MMMB [19]/HITON [20]/PCMB [21] [22]/STMB [23]	两阶段分治；交错执行或对称性校验。	解耦子任务提升了精度；对称校验(AND/OR)增强了理论保障。	严格校验导致召回率低；多轮迭代计算开销大。
FCI [4] [12]/RFCI [13]/F ² CI [24]	引入 PAG/MAG 图模型；推广 d-分离识别潜在混杂及部分祖先关系。	可处理潜变量与选择偏差；RFCI 等改进算法缓解计算复杂度问题。	结果仍存在不确定性；部分定向规则保守。

总体来看，约束型方法在理论上具有良好的可解释性，其结构推断过程能够明确对应观测到的条件独立性证据，从而便于追溯特定边或方向决策的统计依据。此外，该类方法不依赖于特定的全局参数化模型，在条件独立性检验可靠且相关假设成立的情况下，可提供一致性保证。然而，其性能高度依赖于 CI 检验的准确性。在小样本、高维、变量高度相关或存在非线性与混合分布的情形下，检验效能往往下降，可能导致学习得到的图结构不稳定。

3.2. 基于评分的因果发现

评分型(score-based)因果发现方法将结构学习问题转化为在候选图空间中优化评分函数的任务，通过最大化或最小化该评分来确定最优图结构[25]。与约束型方法关注局部条件独立性约束不同，评分型方法侧重在全局层面平衡模型拟合优度与复杂性，并通过在离散结构空间的搜索实现最优因果图的识别。在贝叶斯网络等常用图模型中，评分函数通常具有可分解性，即每个节点的得分仅依赖于其父节点集合，这为局部更新与高效搜索提供了可能。同时，引入正则化或惩罚项可有效控制模型复杂度，防止过拟合，并在一定假设条件下保证结构学习结果的一致性[26]。

在具体实现上，评分型方法典型采用贪心搜索或局部爬山(Greedy Hill-Climbing)，通过迭代执行边的

添加、删除或反向操作来优化评分。这种策略在计算上具有可扩展性，但容易陷入局部最优，因此通常结合随机重启或禁忌搜索等方法以增强全局探索能力[26]。另一类方法，如等价类搜索(GES)，直接在 DAG 等价类空间中进行优化，从而减少冗余结构的遍历，并在特定假设下保证渐近一致性和理论可保障性。近年来出现的连续优化方法则将 DAG 结构的离散约束转化为可微约束，通过无环性连续刻画将结构学习转化为约束优化问题。现对部分代表性约束方法进行了系统归纳，其分类及优劣势对比如表 2 所示。

Table 2. Score-based causal representative discovery algorithms: classification, core mechanisms, advantages and disadvantages

表 2. 基于评分的因果代表发现算法分类、核心机制及优劣势对比

代表算法	核心机制	优势	局限性
GES [27] [28]	在马尔可夫等价类(CPDAG)空间进行前向加边/后向减边。	突破 DAG 局限；避免局部最优。	计算随维度指数增长；无法识别 V 结构外的方向。
FGES [29]/SGES [30]/GGSL [31]	并行化计算；自适应缓存；图属性剪枝；局部图生长机制。	将评分搜索扩展至百万级变量；大幅提升运算速度。	仍基于离散空间，对非线性关系拟合能力有限。
IMaGES [32]/GIES [33]	独立评分加权聚合；干预马尔可夫等价类搜索。	多源异构偏差校正与基于干预的等价性破除。	对干预机制有特定假设；计算代价依然较高。
GraN-DAG [34]/NOTEARS [35]	神经 SCM 拟合、连续无环约束与梯度优化。	非线性连续优化与可识别性增强。	高计算代价与非凸局部最优风险。

评分型方法的核心优势在于其全局性与高度灵活性：该类方法能够同时整合多种噪声模型、损失函数及正则化策略，并便于与回归、广义线性模型或神经网络等预测建模工具结合，从而实现“结构 + 参数”的统一估计框架[15]。然而，其局限性也十分显著。首先，随着变量数量增加，DAG 空间呈指数增长，使得精确的全局优化在实际应用中难以实现，因此通常依赖启发式或近似搜索策略。其次，评分函数通常隐含特定的分布或函数形式假设，当这些假设与真实数据不符时，可能导致学习结构出现偏差。最后，即便达到全局最优，传统评分型方法仍可能仅识别等价类，无法完全确定所有因果方向，此时需额外信息(如时间序列、干预实验或多环境数据)以进一步提升可识别性。

3.3. 基于混合的因果发现

混合型(hybrid)因果发现方法旨在结合约束型与评分型策略的优势，在统计可靠性与计算效率之间实现平衡。其核心思路是利用条件独立性信息对候选结构空间进行筛选或约束，同时借助评分函数对剩余候选图进行全局优化与评估[36]。相比单一方法，混合策略的提出主要源于两方面考虑：一方面，纯依赖条件独立性检验在小样本或高维数据下可能因统计功效不足而导致学习结果不稳定；另一方面，仅使用评分函数进行全局搜索则面临指数级的结构空间增长及潜在模型错设的风险。通过在学习过程中同步应用独立性约束和评分准则，混合方法能够有效缩减搜索空间，并在实践中提升结构恢复的稳定性和准确性[36]。

在具体实现层面，混合型方法通常呈现出若干典型流程模式。最常见的是“约束先行、评分精炼”策略：首先通过条件独立性检验确定图结构骨架或候选邻域，从而大幅缩减潜在边集合；随后在此受限空间内应用评分函数进行结构优化与方向判定，以增强全局模型的一致性。另一类策略为“评分驱动、约束剪枝”，即在评分搜索过程中引入条件独立性约束，对部分候选操作进行过滤或校验，从而避免生成明显违反独立性结构的图[37]。此外，混合型方法中还存在“局部结构学习与全局整合”的策略，即先

在每个节点的局部范围内识别潜在父节点或马尔可夫邻域，再通过全局一致化过程协调各局部结构之间的冲突，并确保整体图的无环性。该方法在高维数据场景中尤为适用，因为局部筛选不仅能够有效减少候选边的数量，还可显著缓解全局搜索所带来的计算负担[31]。现对部分代表性混合方法进行了系统归纳，其分类及优劣势对比如表 3 所示。

Table 3. Hybrid causal representative discovery algorithms: classification, core mechanisms, advantages and disadvantages
表 3. 基于混合的因果代表发现算法分类、核心机制及优劣势对比

代表算法	核心机制	优势	局限性
MMHC [19] [25] [36] [38]/ H ² PC [39]/HLCD [40]	CI 检验构建骨架；贪婪评分 搜索优化结构。	结合了约束的高效与评分的 鲁棒；显著缩减搜索空间。	骨架误差易传播；高维一 致性理论不足。
ARGES [41]/ESLH [42]	受限评分搜索；动态骨架修 正策略。	提供高维统计一致性保证； 降低计算骨架开销。	依赖超集搜索空间假设； 评分阶段仍较复杂。
GraN-LCS [43]	神经网络拟合 + 连续无环 性约束；Masked 局部约束。	可建模复杂非线性关系；复 杂度较低	训练成本高；超参数敏 感；可解释性较弱。
DE-DAG [44]	数据增强约束	解决小样本统计不稳问题。	数据依赖初始分布质量

总体来看，混合型方法在高维、小样本或噪声较强的数据环境中通常表现出较好的稳健性：与纯约束型方法相比，其结构学习结果对单次条件独立性检验的误差更不敏感；而与纯评分型方法相比，其候选结构空间显著受限，从而有效提升了计算效率。此外，混合型方法在结构推断的解释性方面通常更为突出：研究者既可依托条件独立性证据说明特定边被排除的统计依据，又可借助评分函数展示最终模型在全局拟合与复杂度之间的平衡。然而，这类方法仍存在若干挑战。首先，在有限样本条件下，条件独立性检验的误差与评分函数偏差可能相互叠加，导致过度剪枝或评分偏倚；其次，算法性能对参数设定较为敏感，包括显著性水平、最大条件集规模、评分惩罚强度及候选边数量等。此外，由于混合方法通常包含多阶段启发式流程，其一致性与可识别性分析较为复杂，理论保障往往依赖于具体实现及所采用的假设条件。

3.4. 三类因果发现范式的多维比较与评析

为更准确把握三类因果发现范式的适用边界及方法局限，现从理论前提、数据适配性、计算与样本代价、稳健性以及结果确定性五个方面，对典型算法进行归纳性比较与批判性分析。

首先，在理论基础方面，三类方法各自依赖不同的前提条件。约束型方法以马尔可夫性和忠实性为主要理论支点，逻辑结构清晰，但在真实数据中易受潜在混杂、近确定性关系等因素影响，从而削弱推断有效性。评分型方法则更多建立在特定参数模型或函数形式之上，当数据生成机制偏离既定假设时，容易出现模型错设。混合型方法虽然试图兼顾两类思路，但其两阶段流程往往伴随误差传递问题，因而一致性保障相对较弱。

其次，从数据适配性及资源开销来看，不同范式表现出明显差异。约束型方法对离散数据及部分连续数据具有较好适应性，但在高维条件下，随着条件集扩大，检验开销迅速增加，小样本情形下尤易出现识别不稳。评分型方法借助连续优化在高维连续数据中展现出较好的计算效率，但对混合型数据和复杂分布的兼容性仍有限，且通常依赖更充足的数据与算力支持。混合型方法通过先筛选后搜索的机制压缩了搜索空间，因此在效率与精度之间往往能够取得较为平衡的表现。

再次，在抗噪声干扰和抗局部错判方面，评分型方法通常具有更好的稳健性，因为其依据全局评分函数进行结构选择，能够在一定程度上削弱局部扰动对整体学习结果的影响。相比之下，约束型方法高度依赖局部条件独立性检验，一旦关键检验失误，错误可能在后续结构恢复中持续放大。不过，当模型假设明显偏离真实机制时，评分型方法的全局最优也可能稳定地导向错误结构，而约束型方法由于参数依赖较弱，在某些错设场景下反而更具稳健性。

最后，就结果明确性而言，多数传统方法只能恢复马尔可夫等价类，而难以唯一确定全部边方向。无论是 CPDAG 还是 PAG，本质上都保留了部分方向不确定性，这在实际应用中会限制结果的直接解释力。

综上，三类因果发现范式在理论支撑、数据适配、计算代价、稳健性与结果解释性等方面各有优势与边界。为更直观呈现上述比较结论，表 4 将进一步对不同范式及其代表性算法的综合表现进行整理与归纳。

Table 4. Multi-dimensional comparative analysis of representative causal discovery algorithms

表 4. 代表性因果发现算法多维度横向比较与批判性分析

算法范式	代表算法	理论假设 依赖度	数据类型 适应性	计算/样本效率	对噪声/错设的 鲁棒性	输出确定性与 解释力
基于约束	PC [11] FCI [4] [12]	强依赖忠实性 与 CI 检验准确 性	擅长离散/线性 连续，非线性 适应差	高维易指数爆 炸，小样本极 不稳定	局部检验误差 易发生级联传 播	输出 CPDAG/PAG， 方向不确定性 高，但推断过 程透明
基于评分	GES [27] [28] NOTEARS [35]	强依赖参数化 模型与分布假 设	连续优化擅长 连续数据，混 合数据处理弱	连续优化提升 了高维效率， 但需大样本	抗局部噪声 强，但极易受 模型错设影响	部分方法可打 破等价类，但 深度模型解释 性较弱
基于混合	MMHC [19] [25] [36] [38] GraN-LCS [43]	假设依赖介于 两者之间	通过灵活组合 可适应多种数 据类型	约束剪枝大幅 降低搜索空 间，效率最优	综合两者优 势，对单次检 验误差 insensitive	输出多为等价 类，解释性兼 具统计与拟合 双重依据

4. 总结

本文聚焦于观测数据条件下的因果发现问题，对现有方法体系进行了系统化梳理与比较。以图模型语言，特别是有向无环图及其等价类表示作为分析框架，强调因果发现不仅是结构拟合，更是一个由可识别性边界、统计证据和计算可扩展性共同约束的推断任务。

在方法论层面，以“约束型 - 评分型 - 混合型”三分法作为主轴，力求在统一坐标系下梳理不同路线的共性与差异。约束型方法以条件独立关系为主要信息来源，推断链条透明、便于解释，但其性能高度依赖独立性检验的可靠性，并在有限样本、高维与复杂分布情形下面临显著不稳定性；评分型方法将结构学习表述为目标函数优化问题，能够自然整合似然、正则化与多样建模器，并与连续优化、可微结构学习等现代计算范式形成紧密衔接，但也不可避免地受到搜索空间巨大与模型错误风险的制约；混合型方法通过在约束信息与得分准则之间建立互补机制，往往在实践中展现更好的效率与稳健性，但其理论分析与一致性保证更依赖具体流程设计与假设条件。总体而言，这三类范式共同刻画了因果发现研究在“统计证据 - 可识别性 - 计算策略”三者之间的权衡结构。

5. 挑战与未来方向

纵观因果发现领域的演进历程，尽管相关算法在理论基础与计算架构层面已取得显著突破，但当其被部署于真实世界高度复杂的数据场景时，依然暴露出若干亟待解决的瓶颈问题。对这些核心局限性进行系统剖析，不仅能够清晰界定现有方法的适用范围，也为后续技术范式的迭代提供了关键指引。

首先，未观测混杂因子(Unobserved Confounders)与潜变量干扰始终是阻碍因果推断置信度提升的首要难题。以 FCI 为代表的现有框架虽然尝试借助局部祖先图(PAG)来表征部分隐变量的作用机制，但其推断结果往往充斥着大量无法确定方向的边缘，难以输出清晰的因果拓扑。更棘手的是，一旦潜变量与可观测特征之间存在复杂的非线性映射，那些依赖线性或高斯分布假设的传统识别准则便会面临失效风险。针对这一困境，未来的研究重心或将转向基于深度生成模型(例如变分自编码器 VAE 及生成对抗网络 GAN)的潜变量表征学习。通过引入更具约束力的数据分布先验，抑或融合多环境下的干预数据，有望实现对未观测混杂因子的有效解耦与可识别性重构。

其次，数据分布的异质性与非平稳特征(Non-stationarity)对因果结构的鲁棒性构成了严峻挑战。绝大多数经典算法均以独立同分布作为前提假设，然而在医疗多中心研究、跨域分析或长周期时序观测等真实场景中，数据往往伴随着显著的分布漂移或底层机制变异。如何在动态变化的环境中精准剥离“波动的边缘分布”与“恒定的因果机制”，是当前学术界攻坚的焦点。未来的技术演进可着眼于因果不变性学习(Causal Invariance Learning)与独立机制原则(Independent Mechanism Principle)，将不同时间切片或环境域下的分布差异视为一种“自然干预”信号，借此反向约束并推演具备跨域泛化能力的因果骨架。

再者，超高维变量空间(High-dimensionality)引发的计算爆炸与统计效能衰减问题依然严峻。在基因组学或全脑网络分析等前沿领域，变量规模动辄突破万级甚至百万级大关。在此背景下，基于条件独立性检验的约束型算法极易陷入“维度灾难”并伴随严重的假阳性激增；而评分型方法所面临的结构搜索空间更是呈现超指数级膨胀。尽管以 NOTEARS 为代表的连续优化范式成功将离散图搜索转化为可微约束问题，但其在非凸优化过程中易陷入局部最优的缺陷，以及向非线性场景扩展时高昂的算力开销，依然是不容忽视的短板。破局之道可能在于两个维度：其一，深度融合大语言模型(LLM)或垂直领域知识图谱，通过引入强先验知识大幅修剪无效的搜索分支；其二，探索基于图神经网络(GNN)与强化学习的分布式因果发现新架构，力求在超大规模网络中实现线性时间复杂度的近似推断。

最后，因果发现算法的评估体系与模型可解释性亟待重塑。目前，主流算法的性能验证过度依赖于人工合成数据集，而在缺乏真实因果基准(Ground Truth)的实际业务场景中，往往难以对其有效性进行客观评判。未来的工作需要致力于构建更贴近真实世界复杂度的基准测试平台(Benchmark)，并积极探索将因果发现模块与下游决策任务(如反事实推理、强化学习)深度耦合的端到端评估机制。与此同时，进一步增强深度因果模型(如神经结构因果模型 SCM)的内在透明度，确保其推断逻辑能够被领域专家直观理解与严格校验，这将是推动因果发现技术从实验室走向高风险工业及医疗应用场景的必由之路。

致 谢

作者祁慧英衷心感谢以上各位作者在研究过程中给予的宝贵指导和深刻建议，同时感谢同门之间的建设性讨论和技术协助。此外，作者还感谢匿名评审专家提出的有益意见和建议。

参考文献

- [1] McDonald, R.P. (2002) Judea Pearl. Causality: Models, Reasoning, and Inference. Cambridge University Press. <https://doi.org/10.1007/bf02294851>
- [2] Jonas, P., Janzing, D. and Scholkopf, B. (2017) Elements of Causal Inference: Foundations and Learning Algorithms. MIT Press.

-
- [3] Sewall, W. (1921) Correlation and Causation. *Journal of Agricultural Research*, **20**, Article 557.
- [4] Spirtes, P., Glymour, C. and Scheines, R. (2001) Causation, Prediction, and Search. 2nd Edition, The MIT Press. <https://doi.org/10.7551/mitpress/1754.001.0001>
- [5] Judea, P. (2014) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Elsevier.
- [6] Daphne, K. and Friedman, N. (2009) Probabilistic Graphical Models: Principles and Techniques. MIT Press.
- [7] Judea, P. (2009) Causality. Cambridge University Press.
- [8] Ioannis, T. and Aliferis, C.F. (2003) Towards Principled Feature Selection: Relevancy, Filters and Wrappers. *International Workshop on Artificial Intelligence and Statistics*, Florida, 3-6 January 2003.
- [9] Spirtes, P. (2010) Introduction to Causal Inference. *Journal of Machine Learning Research*, **11**, 1643-1662.
- [10] Verma, T. and Pearl, J. (2022) Equivalence and Synthesis of Causal Models. In: *Probabilistic and Causal Inference*, ACM, 221-236. <https://doi.org/10.1145/3501714.3501732>
- [11] Spirtes, P. and Glymour, C. (1991) An Algorithm for Fast Recovery of Sparse Causal Graphs. *Social Science Computer Review*, **9**, 62-72. <https://doi.org/10.1177/089443939100900106>
- [12] Spirtes, P.L., Meek, C. and Richardson, T.S. (2013) Causal Inference in the Presence of Latent Variables and Selection Bias.
- [13] Colombo, D., Maathuis, M.H., Kalisch, M. and Richardson, T.S. (2012) Learning High-Dimensional Directed Acyclic Graphs with Latent and Selection Variables. *The Annals of Statistics*, **40**, 294-321. <https://doi.org/10.1214/11-aos940>
- [14] Judea, P. and Verma, T.S. (1995) A Theory of Inferred Causation. In: *Studies in Logic and the Foundations of Mathematics. Vol. 134*, Elsevier, 789-811.
- [15] Tsamardinos, I., Aliferis, C.F. and Statnikov, A. (2003) Algorithms for Large Scale Markov Blanket Discovery. American Association for Artificial Intelligence.
- [16] Yaramakala, S. and Margaritis, D. (2005) Speculative Markov Blanket Discovery for Optimal Feature Selection. *15th IEEE International Conference on Data Mining (ICDM'05)*, Houston, 27-30 November 2005, 1-4.
- [17] Yang, X., Wang, Y., Ou, Y. and Tong, Y. (2019) Three-Fast-Inter Incremental Association Markov Blanket Learning Algorithm. *Pattern Recognition Letters*, **122**, 73-78. <https://doi.org/10.1016/j.patrec.2019.02.002>
- [18] Margaritis, D. (2009) Toward Provably Correct Feature Selection in Arbitrary Domains. *Advances in Neural Information Processing Systems*, **22**, 1-9.
- [19] Tsamardinos, I., Aliferis, C.F. and Statnikov, A. (2003) Time and Sample Efficient Discovery of Markov Blankets and Direct Causal Relations. *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, 24-27 August 2003, 673-678. <https://doi.org/10.1145/956750.956838>
- [20] Aliferis, C.F., Tsamardinos, I. and Statnikov, A. (2003) HITON: A Novel Markov Blanket Algorithm for Optimal Variable Selection. *AMIA Annual Symposium Proceedings*, Washington, 8-12 November 2003, 21-25.
- [21] Peña, J.M., Björkegren, J. and Tegnér, J. (2005) Scalable, Efficient and Correct Learning of Markov Boundaries under the Faithfulness Assumption. In: *Lecture Notes in Computer Science*, Springer, 136-147. https://doi.org/10.1007/11518655_13
- [22] Peña, J.M., Nilsson, R., Björkegren, J. and Tegnér, J. (2007) Towards Scalable and Data Efficient Learning of Markov Boundaries. *International Journal of Approximate Reasoning*, **45**, 211-232. <https://doi.org/10.1016/j.ijar.2006.06.008>
- [23] Gao, T. and Ji, Q. (2016) Efficient Markov Blanket Discovery and Its Application. *IEEE Transactions on Cybernetics*, **47**, 1169-1179. <https://doi.org/10.1109/tcyb.2016.2539338>
- [24] Strobl, E.V. (2017) Causal Discovery under Non-Stationary Feedback. University of Pittsburgh.
- [25] Heckerman, D., Geiger, D. and Chickering, D.M. (1995) Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *Machine Learning*, **20**, 197-243. <https://doi.org/10.1023/a:1022623210503>
- [26] Heckerman, D. (1998) A Tutorial on Learning with Bayesian Networks. In: *Learning in Graphical Models*, Springer, 301-354. https://doi.org/10.1007/978-94-011-5014-9_11
- [27] Chickering, D.M. (2002) Optimal Structure Identification with Greedy Search. *Journal of Machine Learning Research*, **3**, 507-554.
- [28] Aragam, B. (2024) Greedy Equivalence Search for Nonparametric Graphical Models. arXiv:2406.17228.
- [29] Ramsey, J., Glymour, M., Sanchez-Romero, R. and Glymour, C. (2017) A Million Variables and More: The Fast Greedy Equivalence Search Algorithm for Learning High-Dimensional Graphical Causal Models, with an Application to Functional Magnetic Resonance Images. *International Journal of Data Science and Analytics*, **3**, 121-129. <https://doi.org/10.1007/s41060-016-0032-z>
- [30] Chickering, D.M. and Meek, C. (2015) Selective Greedy Equivalence Search: Finding Optimal Bayesian Networks Using

a Polynomial Number of Score Evaluations.

- [31] Gao, T., Fadnis, K. and Campbell, M. (2017) Local-to-Global Bayesian Network Structure Learning. *International Conference on Machine Learning*, Sydney, 6-11 August 2017.
- [32] Ramsey, J.D., Hanson, S.J., Hanson, C., Halchenko, Y.O., Poldrack, R.A. and Glymour, C. (2010) Six Problems for Causal Inference from fMRI. *NeuroImage*, **49**, 1545-1558. <https://doi.org/10.1016/j.neuroimage.2009.08.065>
- [33] Hauser, A. and Bühlmann, P. (2012) Characterization and Greedy Learning of Interventional Markov Equivalence Classes of Directed Acyclic Graphs. *The Journal of Machine Learning Research*, **13**, 2409-2464.
- [34] Lachapelle, S., Brouillard, P., Deleu, T., *et al.* (2019) Gradient-Based Neural DAG Learning.
- [35] Zheng, X., Aragam, B., Ravikumar, P., *et al.* (2018) DAGs with No Tears: Continuous Optimization for Structure Learning. *Advances in Neural Information Processing Systems*, **31**, 1-22.
- [36] Tsamardinos, I., Brown, L.E. and Aliferis, C.F. (2006) The Max-Min Hill-Climbing Bayesian Network Structure Learning Algorithm. *Machine Learning*, **65**, 31-78. <https://doi.org/10.1007/s10994-006-6889-7>
- [37] Scanagatta, M., Salmerón, A. and Stella, F. (2019) A Survey on Bayesian Network Structure Learning from Data. *Progress in Artificial Intelligence*, **8**, 425-439. <https://doi.org/10.1007/s13748-019-00194-y>
- [38] Scutari, M. (2010) Learning Bayesian Networks with Thebnlearnrpackage. *Journal of Statistical Software*, **35**, 1-22. <https://doi.org/10.18637/jss.v035.i03>
- [39] Gasse, M., Aussem, A. and Elghazel, H. (2014) A Hybrid Algorithm for Bayesian Network Structure Learning with Application to Multi-Label Learning. *Expert Systems with Applications*, **41**, 6755-6772. <https://doi.org/10.1016/j.eswa.2014.04.032>
- [40] Ling, Z.L., Peng, H.H., Zhang, Y.W., *et al.* (2024) Hybrid Local Causal Discovery. arXiv:2412.19507.
- [41] Nandy, P., Hauser, A. and Maathuis, M.H. (2018) High-Dimensional Consistency in Score-Based and Hybrid Structure Learning. *The Annals of Statistics*, **46**, 3151-3183. <https://doi.org/10.1214/17-aos1654>
- [42] Wang, N., Liu, H., Zhang, L., Cai, Y. and Shi, Q. (2024) An Efficient Skeleton Learning Approach-Based Hybrid Algorithm for Identifying Bayesian Network Structure. *Engineering Applications of Artificial Intelligence*, **133**, Article 108105. <https://doi.org/10.1016/j.engappai.2024.108105>
- [43] Liang, J., Wang, J., Yu, G., Domeniconi, C., Zhang, X. and Guo, M. (2023) Gradient-Based Local Causal Structure Learning. *IEEE Transactions on Cybernetics*, **54**, 486-495. <https://doi.org/10.1109/tcyb.2023.3237635>
- [44] Huang, X., Guo, X., Li, Y. and Yu, K. (2023) A Novel Data Enhancement Approach to DAG Learning with Small Data Samples. *Applied Intelligence*, **53**, 27589-27607. <https://doi.org/10.1007/s10489-023-04999-2>