

19. Daugavietis J.E., Soloha R., Dace E., Ziemele J.A. Comparison of Multi-Criteria Decision Analysis Methods for Sustainability Assessment of District Heating Systems, *Energies*, 2022, 15, 2411.
20. Kizielewicz B., Bączkiewicz A. Comparison of Fuzzy TOPSIS, Fuzzy VIKOR, Fuzzy WASPAS and Fuzzy MMOORA methods in the housing selection problem, *Procedia Computer Science*, 2021, 192, pp. 4578-91.
21. Salabun W., Wątróbski J., Shekhovtsov A. Are MCDA Methods Benchmarkable? A Comparative Study of TOPSIS, VIKOR, COPRAS, and PROMETHEE II Methods, *Symmetry*, 2020, 12, pp. 1549.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

Подоплелова Елизавета Сергеевна – Южный федеральный университет; e-mail: chuzhinova@sfnu.ru; г. Таганрог, Россия; тел.: +79525844188; аспирант.

Podoplelova Elizaveta Sergeevna – Southern Federal University; e-mail: chuzhinova@sfnu.ru; Taganrog, Russia; phone: +79525844188; postgraduate student.

УДК 519.254

DOI 10.18522/2311-3103-2023-3-125-134

М.Ю. Георги

МЕТОДЫ ИЗВЛЕЧЕНИЯ ПРИЧИННОСТИ ИЗ ДАННЫХ НАБЛЮДЕНИЙ В ПРАКТИКЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Обсуждается значимость извлечения причинно-следственных связей в машинном обучении для принятия решений и оценки воздействия на реальный мир. Отмечается, что большинство текущих успехов в машинном обучении основаны на огрубленном распознавании образов и корреляционном анализе, но для более сложных задач необходимо извлекать причинно-следственные связи. Декларируется что проблемы объяснимости прогнозов и причинно-следственного понимания, даже с применением передовых методик машинного обучения LIME, SHAP, TreeSHAP, DeepSHAP, Shapley Flow, являются фундаментальными препятствиями в развитии искусственного интеллекта. В статье кратко раскрываются основные философские математические концепции и определения причинности, включая понятия контрфактуалов, байесовских сетей, направленных ациклических графов и причинно-следственного формального вывода. Делается вывод о том, что практическая значимость базирующегося на данных причинно-следственного анализа, состоит в ответах на априори сформулированные вопросы, которые могут отражать гипотетическую связь между событием (причиной) и вторым событием (следствием), где второе событие является прямым следствием первого. Далее производится сравнительный анализ способов и основных сценариев использования фреймворков Causal Discovery и Causal Inference, на базе которых возникает возможность сделать предположения о расположенной в основе исследуемого набора данных причинно-следственной структуре и задействовать статистические методы для оценки силы и направления таких связей. В статье также обсуждаются методы и алгоритмы причинно-следственного анализа и их применение в реальных задачах. Упомянуты репрезентативные методы, такие как модели на основе ограничений, модели на основе оценок и функциональные каузальные модели, тесты (условные) на независимость, оценочные функции, всё то что может быть задействовано для решения проблемы извлечения причинно-следственных связей из наблюдательных данных, большинство которых реализовано в open-source фреймворки, таких как Microsoft DoWhy, Uber CausalML, causal-learn, Econ-ml и многие другие, которые реализуют большинство упомянутых методов причинно-следственного анализа.

Причинно-следственный ИИ; Объективность ИИ; скоростное обучение; прогноз под вмешательством.

M.Y. Georgi

METHODS FOR MINING CAUSUSITY FROM OBSERVATIONS IN ARTIFICIAL INTELLIGENCE

The article discusses the importance of capturing causal relationships in machine learning for decision-making and evaluating real-world impact. It is noted that most current successes in machine learning are based on pattern recognition and correlation analysis; however, for more complex tasks, extracting causal relationships is necessary. The problems of explainability of predictions and causal understanding, even with the use of advanced machine learning techniques such as LIME, SHAP, TreeSHAP, DeepSHAP, and Shapley Flow, are recognized as fundamental obstacles in the development of artificial intelligence. The article briefly presents the main philosophical and mathematical concepts and definitions of causality, including counterfactuals, Bayesian networks, directed acyclic graphs, and causal formal inference. It concludes that the practical significance of data-based causal analysis consists in answering a priori formulated questions, which may reflect a hypothetical relationship between an event (a cause) and a second event (an effect), where the second event is a direct consequence of the first. A comparative analysis of the methods and main scenarios for using the Causal Discovery and Causal Inference frameworks is also carried out. Based on this analysis, it becomes possible to make assumptions about the causal structure underlying the investigated dataset and to use statistical methods to evaluate the strength and direction of such relationships. The article also discusses methods and algorithms of causal analysis and their application in real-world tasks. Representative methods are mentioned, such as constraint-based models, estimation-based models, functional causal models, (conditional) independence tests, evaluation functions, and other tools that can be used to solve the problem of extracting causal relationships from observational data. Most of these methods are implemented in open-source frameworks such as Microsoft DoWhy, Uber CausalML, causal-learn, Econ-ml, and many others, which facilitate causal analysis.

Causal AI; AI fairness; few-shot fine-tuning; counterfactual reasoning.

1. Введение. Необходимость причинно-следственного вывода. Машинное обучение привлекательно тем, что его инструментарий можно применять практически к безграничному кругу задач, от обработки естественного языка до компьютерного зрения и медицинской диагностики. Большинство текущих успехов сводятся к огрубленному распознаванию образов на надлежащим образом собранных независимых и идентично распределенных данных (**iid data**). Поэтому в машинном обучении, в особенности "с учителем", корреляционный анализ долгое время остается фундаментом прогнозирования целевой переменной исходя из переменных признаков. Эта стратегия хорошо зарекомендовала себя на практике, поэтому современные достижения ИИ, такие как авто-регрессионные большие языковые модели AR-LLM (ChatGPT), становятся едва ли не самой важной составляющей технологического прогресса.

Естественно, что подходы, связанные с Большими Данными (и всё более увеличивающимися), особенно в последние 2 десятилетия вывели на первый план мнение, что «корреляция важнее причинно-следственной связи», и даже модели непараметрической статистики прекрасно работают, задействуя только корреляты (условную вероятность): «Причинность не будет отброшена, но она сброшена с пьедестала как основной источник смысла. Большие данные стимулируют некаузальный анализ, часто заменяя причинно-следственные связи» [1].

Джуда Перл и Дэйна Маккензи переформулировали эту концепцию следующим образом: «Надежда – а в настоящее время она пока молчит – заключается в том, что сами данные будут вести нас к правильным ответам всякий раз, когда возникают причинные вопросы» [2].

Однако во многих прикладных областях эти ожидания столкнулись с фундаментальными препятствиями. Проблемы с устойчивостью, с распознаванием и реагированием на новые обстоятельства, для которых модели традиционного ИИ не могут быть явно запрограммированы, или даже обучены, замедляют переход от узкого искусственного интеллекта (ANI) к общему (AGI). Даже впечатляющие успехи трансферного (Transfer), прицельного (Zero-shot) и “быстрого” (Few-Shot) обучения, обучения с подкреплением (Reinforcement Learning), адаптации предметной области (Domain Adaptation) так и не сняли проблему объяснимости прогнозов и причинно-следственного понимания получаемых результатов, являющегося фундаментальным свойством человеческого сознания.

Между формальными причинными и ассоциативными умозаключениями пролегал качественное различие: если вывод о причинности анализирует реакцию переменной-следствия при изменении потенциальной причины, то корреляция указывает всего лишь на общую взаимосвязь [3]. Таким образом, изучение причинно-следственных связей представляет собой задачу повышенной сложности.

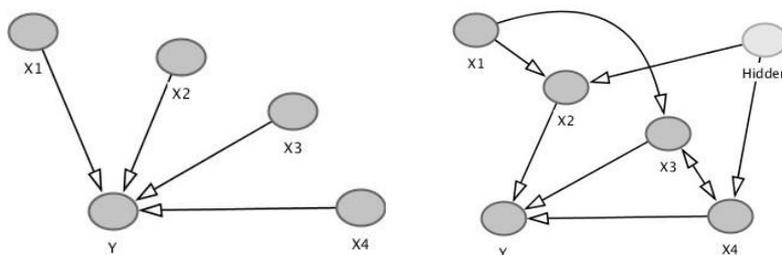


Рис. 1. Постановка традиционной модели машинного обучения (слева) и модели каузальной (справа) [4]. Дополнительно вводятся новые уровни абстракции: общие причины (конфаундеры), общие эффекты (коллайдеры), медиаторы, предусловители и т.д.

Поскольку все более важной целью моделирования становится не прогноз и его относительное качество, но основанный на разрабатываемой модели способ принятия решений и оценка вмешательств в реальный мир, возникла потребность выявлять основные механизмы, почему что-то произошло и количественно оценивать фактическое влияние одной переменной (ковариаты) на цель, реакцию, потенциальный эффект. Методы машинного обучения, ориентированные на сходные задачи, например *LIME*, *SHAP*, *TreeSHAP*, *DeepSHAP*, *Shapley Flow* можно использовать только с уже обученными моделями, что накладывает ограничение на их объяснительную силу [5].

По мнению Джуды Перла, высказанному также в «Книге причин» [2], каузальные модели являются, необходимым, но пока недостаточным элементом в развитии сильного искусственного интеллекта (ASI): «День, когда ИИ сможет приблизиться к человеческому интеллекту, не так далёк, но пока его возможности следует оценивать по трем уровням когнитивных способностей: видение (ассоциация), действие (вмешательство) и воображение (контрфактуальность), формируя так называемую, каузальную лестницу Перла. До недавнего времени искусственный интеллект оперировал лишь на самом нижнем её уровне – видении».

2. Некоторые общие сведения о причинности. Причинность – это способ понимания окружающей среды с точки зрения взаимодействия, то, что можно узнать о мире в результате вмешательств в его предполагаемые механизмы, при этом наблюдение моделируется как их побочный эффект. Установление причинности, изначально мыслительный процесс, описывающий идеи о природе отношений причины и следствия, проверяет, существует ли связь причины со следствием.

Вероятно, наиболее ранняя попытка разработать строгое определение причинности была предпринята Дэвидом Юмом [6]. Концепции Юма, компактно оформленные в терминах контрфактуалов, определяют причинность следующим образом: «Мы рассматриваем причину как что-то создающее различие, и это различие должно отличаться от того, что произошло бы без него. Если бы его не было, его эффекты – по крайней мере, некоторые из них, а обычно и все – тоже отсутствовали бы» [7].

Более формальное определение принципа общей причинности было предложено философом Хансом Райхенбахом. Если две случайные величины X и Y статистически зависимы ($X | Y$), то либо (a) X вызывает Y , (b) Y вызывает X , либо (c) существует третья переменная Z , которая вызывает как X , так и Y . Далее X и Y становятся независимыми при обусловливание (*conditioning*) на Z , т.е. $X \perp Y | Z$. (*Обуславливание переменной – это механизм формирования условий для проведения гипотетических экспериментов, включающий в себя анализ значений других переменных для данного значения обусловленной переменной.*). В этом случае говорят, что Z «экранирует» зависимость между X и Y [8].

Данная концепция легла в основу байесовских графовых моделей (причинно-следственных вероятностных сетей BBN). Задействуя такой инструментарий, причинный формальный вывод (**causal inference**) стало возможным с помощью данных пассивных наблюдений, при условии их достаточности, “зафиксировать” в модели сопутствующие переменные (control on confounders) и строго оценить эффект внешнего вмешательства на получаемый результат. Методы причинно-следственного анализа (ПСА) используют объемные, глубокие предположения и соглашения, которые позволяют оценить контрфактический (ненаблюдаемый) результат, отвечая на вопросы 3-го уровня лестницы Перла “Почему?” [9].

Таким образом ответы на причинные вопросы, выдвижение-подтверждение-опровержение гипотез о механизме порождения данных, лежащих в основе исследуемого феномена, собственно, и являются основной сферой применения ПСА. Подобные вопросы особенно важны в сценариях принятия решений, связанных с оценкой гипотетической ретроспекции (контрфакта):

- ◆ Улучшит ли показатели предлагаемое изменение в системе? (*Будет ли это работать*).
- ◆ Что привело к изменению показателей системы? (*Почему это сработало*)
- ◆ Какие изменения в системе могут улучшить результаты для пользователей? (*Что необходимо предпринять*).
- ◆ Как система взаимодействует с поведением человека и каково влияние, например системы рекомендаций, на активность людей? (*Каковы общие эффекты/последствия*).

3. Основные понятия и определения ПСА. На практике исследование причинности проводится по двум связанным, но относящимся к разным аспектам процесса понимания причинности в данных, концепциям: **Выявление Причинности** (Causal Discovery), **Оценка Причинности** (Causal Inference/Estimation) [10]. Первое направление решает задачу раскрытия взаимосвязей между переменными в наборе данных, не обязательно делая какие-либо предположения о направлении или силе причинно-следственных связей, в то время как второе сосредоточено на численной оценке причинно-следственного эффекта вмешательства на результат.

Таблица 1

Сравнительный анализ сценариев использования методологий выявления причинности (ВП) и оценки причинности (ОП)

Выявление Причинности (ВП)	Оценка Причинности (ОП)
<u>Занимается</u> выявлением лежащей в основе системы причинно-следственной структуры на основе данных наблюдений.	<u>Занимается</u> определением влияния одной переменной на другую с учетом известной причинно-следственной структуры.
<u>Использует</u> методы теории графов, таких как байесовские сети, марковские сети и моделирование структурных уравнений.	<u>Использует</u> методы статистики и машинного обучения, такие как линейная регрессия, логистическая регрессия и нейронные сети.
<u>Фокусируется</u> на обнаружении причинно-следственных связей между переменными.	<u>Фокусируется</u> на определении силы причинно-следственных отношений и составлении прогнозов на их основе.
<u>Важно</u> для понимания того, как работает система, и прогнозирования того, как изменения в одной переменной повлияют на изменения в других переменных.	<u>Важно</u> для разработки вмешательств, которые улучшат производительность системы.
Причинно-следственные связи могут быть выведены из моделей статистической зависимости и независимости между переменными.	Причинно-следственные связи могут быть использованы для прогнозирования влияния вмешательств на переменные в системе.
Методы причинно-следственного обнаружения включают графовые модели, теоретико-информационные подходы и байесовские сети.	Методы причинно-следственного вывода включают рандомизированные контролируемые испытания, регрессионный анализ, сопоставление показателей склонности и анализ инструментальных переменных.
<u>Не может</u> установить причинно-следственную связь окончательно, т.к. могут присутствовать ненаблюдаемые вмешивающиеся факторы или обратная причинно-следственная связь.	<u>Может</u> установить причинно-следственную связь более определенно, но только для конкретного изучаемого воздействия (лечения) и в конкретных условиях исследования.
<u>Применяется</u> для обнаружения причинно-следственных связей в сложных системах, где базовая причинно-следственная структура неизвестна.	<u>Применяется</u> для прогнозирования последствий вмешательств в сложных системах.

3.1. Методология выявления причинности (ВП). Поиск зависимостей, иллюстрирующих отношения, присущие данным, опирается на фундаментальное свойство Acyclicity (Ацикличность) и сопровождается **автоматизированным** построением наиболее правдоподобного направленного ациклического графа (DAG). Для данного графа $\mathcal{G} = (\mathbf{X}, \mathcal{E})$ случайные переменные \mathbf{X} образуют Байесову сеть в отношении \mathcal{G} если выполняется формула факторизации, выражающей совместное распределение вероятностей [11].

$$P(\mathbf{X}) = \prod_{X \in \mathbf{X}} P(X | Pa(X)). \quad (1)$$

Иными словами, это топологическая сортировка, без циклов соединяющих другие ребра, где каждый узел находится в определенном порядке. Для реализации *Causal Discovery* применяются методы сканирования наборов данных на предмет условной независимости между переменными, прогонки через набор совместимых моделей, обеспечивающих наблюдаемое распределение, и затем выбора наиболее подходящей, или, если повезет, единственной модели, которая теоретически способна воспроизвести это наблюдаемое распределение. Наиболее подробно методология изложена в [12].

3.2. Алгоритмы выявления причинности (ВП). Репрезентативные методы в первом приближении могут делиться на 4 категории, а именно модели на основе ограничений [13], модели на основе оценок [14], функциональные каузальные модели. Большинство упомянутых методов реализовано в фреймворках DoWhy, CausalML, causal-learn, Econ-ML, CausalImpact и т.д.

Методы, основанные на ограничениях, используют тесты условной независимости между переменными. (PC, Fast Causal Inference (FCI), Inductive Causation (IC)).

Методы, основанные на оценках, используют функцию оценки для ранжирования возможных причинно-следственных графов, рассчитываемую на основе качества соответствия графа данным. (Greedy Equivalence Search (GES), Greedy Interventional Equivalence Search (GIES), Concave Penalized Coordinate Descent with Reparametrization (CCDr)).

Гибридные методы сочетают в себе элементы методов, основанных на ограничениях и оценках. (Structural Agnostic Modeling (SAM), Causal Additive Model (CAM), Causal Generative Neural Network (CGNN), ММНС

Нелинейные и задействующие асимметрию. (Linear Non-Gaussian Acyclic Model (LiNGAM), Nonlinear Additive Noise Model, Post-nonlinear Causal Model (PNL), Granger Causality)

Получаемые на выходе результаты представляют собой структурно-причинные модели (SCM) [15], или, взаимозаменяемо, направленные графовые структурные причинные модели (DGCM), где вводятся дополнительные уровни математического формализма, в виде функциональных отношений между узлами-предками и потомками на (DAG). Этот класс моделей выражается в виде «модели структурных уравнений» (SEM) или функциональной причинно-следственной модели (FCM).

В отличие от систем уравнений (в общем случае дифференциальных), которые тоже могут устанавливать (ПСС), DGCM не требуют начальных (за исключением временных рядов) и граничных условий. Предполагается, что значение каждой переменной на графе является детерминированной функцией ее прямых причин и неизмеренных возмущений (Марковское ограждение). Обычно используется линейная постановка, хотя форма интерполирующей функции может быть произвольной. Также утверждается, что если орграф причинный, и X является причиной Y ($X \rightarrow Y$), то существуют экзогенное вмешательство T , которое напрямую изменяет распределение (или значение) Y , при фиксации, на некоторых значениях, остальных переменных (ковариат) [16].

3.3. Методология оценки причинности. Производится изучение возможных последствий изменения данной системы там, где зависимость уже, предположительно, установлена.

Основной методологией для исчисления ПСС является фреймворк Возможных (потенциальных) исходов Неймана-Рубина (Potential Outcome) [17, 18], в задачу которого входит оценка разности между наблюдаемым состоянием *объекта выборки (unit)*, которому было вменено какое-либо интервенционное воздействие, и этим же юнитом, на которого подобное воздействие осуществлено не было. **Очевидно, что один из этих исходов всегда будет гипотетически ретроспективным, т. е. ненаблюдаемым.**

ГРАФОВАЯ МОДЕЛЬ (CAUSAL DIAGRAMS) – Используя направленные ациклические графы, выражает причинно-следственные связи с точки зрения логической последовательности, а также качественные предположения, простые для понимания. Ограничивается ациклическими причинно-следственными связями. Не может обрабатывать петли обратной связи, и испытывает потенциальные трудности для количественного анализа.

ПОТЕНЦИАЛЬНЫХ ИСХОДОВ (COUNTERFACTUAL) – Выражает причинно-следственные связи в виде разницы потенциальных исходов при различных условиях воздействия. Делает количественные предположения об реакциях юнитов в исследуемой популяции на воздействие. Предполагается отсутствие интерференции между оцениваемыми объектами групп и отсутствие неизмеренных конфаундеров. Причинность может быть определена ошибочно.

STRUCTURAL EQUATIONS MODEL (SEM) – Представляет причинно-следственные связи как систему уравнений, описывающих взаимосвязь переменных друг с другом. Требуются строгие предположения о функциональной форме отношений между переменными. Неточное количественное представление причинно-следственных связей.

SUFFICIENT COMPONENT CAUSE MODEL (SCC) – Представляет причинно-следственные связи как совокупность необходимых и достаточных условий для возникновения следствия. Возник (в основном в теории) в эпидемиологии, требующей уточнения механизмов внутри отдельных юнитов. Ограничиваются ситуациями, когда причинно-следственные связи хорошо понятны и могут быть представлены как совокупность необходимых и достаточных условий.

3.4. Алгоритмы оценки причинности. Математические методы и численные подходы каузального формального вывода (исчисления ПСС) также включают в себя рандомизированные эксперименты (РКИ), естественные эксперименты, обсервационные исследования, инструментальные переменные, сопоставление показателей склонности, разница в разнице, дизайн разрыва регрессии и синтетические методы контроля [19].

Do-calculus – разработанный Judea Pearl [20, 21] представляет собой набор правил, которые позволяют манипулировать причинно-следственными моделями для оценки причинно-следственных эффектов вмешательств. Анализируется тонкая разница между байесовской статистикой $P(y|x)$ и $P(y|do(x))$, где первое описывает, какие значения Y , вероятно, примет, когда X окажется равным x (т.е. наблюдаемое распределение), а второе описывает значения y , когда для X установлено значение x (т.е. интервенционное распределение). Занимает центральное место в формировании понимания и оценке причинно-следственных эффектов, оценке надежности и устойчивости DAG, SCM и Potential Outcome моделей.

Инструментальные переменные (IV) используются для оценки причинно-следственных эффектов при наличии ненаблюдаемых смешанных переменных или обратной причинно-следственной связи. Заключается в поиске переменной (инструмента), которая коррелирует с воздействием, а не с результатом, за исключением ее влияния на лечение.

Сопоставление показателей склонности (PSM) – “псевдо-рандомизация”, группа не экспериментальных методов сопоставления обработанных и контрольных групп на основе их оценок склонности производится попытка сбалансировать тестовые группы по вмешивающимся факторам.

Разность разностей (DiD) – это квази-экспериментальный метод, который сравнивает изменение результатов между группой лечения и контрольной группой с течением времени. Предполагается, что тестовая и контрольная группа следовали бы параллельным тенденциям в отсутствие воздействия.

Оценка методом разрывного дизайна (RDD) – это квази-экспериментальный метод для условий неприменимости рандомизации, который использует разрыв во вменении воздействия на основе пороговой переменной, являющегося следствием отсечки. Сравниваются результаты чуть выше и ниже порога, предполагая, что субъекты, близкие к порогу, похожи во всех аспектах, кроме воздействия.

Метод синтетического контроля (SCM) – это статистический метод, который используется для оценки эффектов некоторых интервенций или событий на группу, когда у нас есть только один наблюдательный случай. Производится оценка ПСС при наличии ненаблюдаемых переменных-конфаундеров, путем построения синтетической контрольной группы, которая соответствует тестовой группе по ковариатам до воздействия.

4. Теоретические допущения ПСА. Фундамент ПСА опирается на значительное количество априорных убеждений и абстрактно вводимых понятий, которые предоставляют возможность перейти от статистических ассоциаций к причинно-следственным связям, из которых можно выделить 5 отправных, при этом способ их применения для ВП и ОП различается [22].

◆ **Causal sufficiency** (для ВП) предполагается, что все переменные, которые не являются прямыми причинами или следствиями переменной результата, не зависят от переменной результата, учитывая прямые причины переменной результата. Используется для определения является ли переменная каузальной или ей можно в модели пренебречь.

◆ **Causal sufficiency** (для ОП) – предполагается, что все вмешивающиеся переменные (конфаундеры), влияющие как на воздействие, так и на исход, измерены и включены в анализ. Если предположение нарушается, причинно-следственный эффект не может быть точно оценен. Используется для определения того, можно ли оценить причинный эффект на основе данных наблюдений.

◆ **Faithfulness** (для ВП) – предполагается, что наблюдаемые закономерности статистической зависимости между переменными согласуются с причинно-следственной структурой базовой модели. Это означает, что, если две переменные причинно не связаны они не имеют статистической зависимости (при выводе о структуре причинной модели на основе данных наблюдений)

◆ **Faithfulness** (для ОП) предполагается, что наблюдаемые закономерности статистической зависимости между переменными согласуются с причинно-следственной структурой базовой модели. Это означает, что, если две переменные не связаны причинно-следственной связью, они не должны быть статистически зависимыми (при оценке достоверности причинных выводов, сделанных на основе экспериментальных данных или данных наблюдений.).

◆ **Stable unit treatment value assumption (SUTVA)** (для ВП) предполагается, что причинно-следственные связи между переменными стабильны во времени и не меняются из-за вмешательств или других внешних факторов. То есть причинно-следственная структура системы остается неизменной даже при воздействии на отдельные переменные.

◆ **Stable unit treatment value assumption (SUTVA)** (для ОП) предполагается, что потенциальный результат каждого объекта группы (юнита, пациента и т.д.) зависит только от воздействия, вмененного этому юниту, и никакому другому. Таким образом между блоками нет помех, и каждый блок считается независимым от других.

◆ **Ignorability/Exchangeability/unconfoundedness** (для ВП) предполагается, что не существует не измеряемых факторов, влияющих как на воздействие, так и на переменную результата, что позволяет выявить причинно-следственные связи, не делая никаких предположений о лежащей в их основе причинно-следственной структуре.

◆ **Ignorability/Exchangeability/unconfoundedness** (для ОП) – предполагает, что вменение воздействия не зависит от потенциальных результатов с учетом наблюдаемых ковариат, что позволит оценить причинно-следственный эффект от воздействия путем “контроля” потенциальных искажающих факторов.

◆ **Positivity** (для ВП) – предполагается, что каждая переменная в DAG имеет некоторую вариацию, которая не объясняется другими переменными графа. Это допущение необходимо для определения принципиальной возможности получить причинный эффект на основе наблюдаемых данных с использованием графовых методов, таких как do-calculus.

◆ **Positivity** (для ОП) – предполагается, что для каждого значения вмешивающихся факторов существует положительная вероятность получения каждого воздействия. Это допущение необходимо для объективной оценки эффектов воздействий с использованием данных наблюдений.

Приведенный перечень предпосылок не является исчерпывающим. Их формулировки иногда разнятся, а применимость может зависеть от конкретного контекста и методологии, используемой для причинного обнаружения и причинного исследования. Хотя эти предположения помогают сузить число возможных моделей, проблему они решают не полностью.

Заключение. Причинно-следственный анализ на базе данных, полученных без знаний о механизме их порождения, например в маркетинговой атрибуции [23], невероятно сложен как в теории, так и на практике. Он опирается на специфические предположения, которые зачастую трудно обеспечить и защитить, так как нет стопроцентно эффективных проверок устойчивости и чувствительности (robustness checks). Аппарат ориентированных ациклических графов DAG и вводимых поверх дополнительных уровней математического формализма SEM (Pearl SCM), (Rubin PO) хотя и обладает большой репрезентативной силой и свойствами вывода, не решает проблему окончательно, поскольку трудно исполнить ключевое предположение о том, что DAG, выбранный среди множества альтернативных, является правильным, позволяя увериться, что мы не изучаем что-то априори неверное. Однако эти соображения не закрывают тему дебатов о реализуемости ПСА на данных пассивных наблюдений, а такие инструменты как **Google CausalImpact**, **Microsoft doWhy**, **Uber Causalml**, **Hyundai gCastle** и.т.д демонстрируют потенциальную возможность автоматизировать этот процесс.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Hassani H., Xu H., Mansi G.* Big data and causality, *Annals of Data Science*, 2018, 5, pp. 133-156.
2. *Pearl J., Mackenzie D.* The book of why. Penguin Books, 2019.
3. *Schölkopf B., Locatello F., Bauer S., Ke N.R., Kalchbrenner N., Goyal A., and Bengio Y.* Toward causal representation learning, *Proceedings of the IEEE*, 2021, 109 (5), pp. 612-634.
4. *Naser M.* Causality, causal discovery, and causal inference in structural engineering, *arXiv preprint arXiv:2204.01543*, 2022.
5. *Hernán M., Robins J.* Estimating causal effects from epidemiological data, *Journal of Epidemiology and Community Health*, 2006 Jul., 60 (7), pp. 578-586.
6. *Leshner J.H.H.* Hume's Analysis of "Cause" and the "Two-Definitions" Dispute, *Journal of the History of Philosophy*, 1973, 11 (3), pp. 387-392.
7. *Lewis D.* Causation, *The journal of philosophy*, 1973, 70 (17), pp. 556-567.
8. *Hofer-Szabó G., Rédei M., Szabó L.E.* On Reichenbach's common cause principle and Reichenbach's notion of common cause, *The British Journal for the Philosophy of Science*, 1999, 50 (3), pp. 377-399.
9. *Pearl J.* An Introduction to Causal Inference, *The International Journal of Biostatistics*, 2010, 6 (2).
10. *Glymour M., Pearl J., Jewell N.P.* Causal inference in statistics: A primer. John Wiley & Sons, 2016.
11. *Bareinboim E., Pearl J.* Causal inference and the data-fusion problem, *Proceedings of the National Academy of Sciences*, 2016, 113 (27), pp. 7345-7352.

12. *Pellet J.P.* Effective Causal Analysis: Methods for Structure Learning and Explanations, 2010. (Doctoral dissertation, ETH Zurich).
13. *Triantafyllou S., Tsamardinos I.* Constraint-based causal discovery from multiple interventions over overlapping variable sets, *The Journal of Machine Learning Research*, 2015, 16 (1), pp. 2147-2205.
14. *Glymour C., Zhang K., Spirtes P.* Review of causal discovery methods based on graphical models, *Frontiers in genetics*, 2019, 10, pp. 524.
15. *Nogueira A.R., Pugnana A., Ruggieri S., Pedreschi D., Gama J.* Methods and tools for causal discovery and causal inference, *Wiley interdisciplinary reviews: data mining and knowledge discovery*, 2022, 12 (2), e1449.
16. *Yao L., Chu Z., Li S., Li Y., Gao J., Zhang A.* A survey on causal inference, *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2021, 15 (5), pp. 1-46.
17. *Rubin D.B.* Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of educational Psychology*, 1974, 66 (5), pp. 688.
18. *Imbens G.W., Rubin D.B.* Causal inference in statistics, social, and biomedical sciences. Cambridge University Press, 2015.
19. *Chernozhukov V., Wüthrich K., Zhu Y.* An exact and robust conformal inference method for counterfactual and synthetic controls, *Journal of the American Statistical Association*, 2021, 116 (536), pp. 1849-1864.
20. *Pearl J.* Causal diagrams for empirical research, *Biometrika*, 1995, 82 (4), pp. 669-688.
21. *Pearl Judea.* The do-calculus revisited, *arXiv preprint arXiv:1210.4852*, 2012.
22. *Vonk M., Malekovic N., Bäck T., Kononova A.* Disentangling causality: assumptions in causal discovery and inference, *Artificial Intelligence Review*, 2023, pp. 1-37.
23. *Teinmaa I., Albert J., Goldenberg D.* Uplift modeling: from causal inference to personalization, *In Companion Proceedings of the Web Conference*, 2021.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

Георги Михаил Юрьевич – Южный федеральный университет; e-mail: mgeorgi@sfedu.ru; г. Таганрог, Россия; тел.: +79185035098; аспирант.

Georgi Mikhail Yurievich – Southern Federal University; e-mail: mgeorgi@sfedu.ru; Taganrog, Russia; phone: +79185035098; postgraduate student.

УДК 004.942:519.254:519.876.5:621.396.96:629.7.018 DOI 10.18522/2311-3103-2023-3-134-143

И.Л. Щербов

ИССЛЕДОВАНИЕ МЕТОДОВ ПРОВЕРКИ ЗНАЧИМОСТИ КОЭФФИЦИЕНТОВ СГЛАЖИВАЮЩЕГО ПОЛИНОМА

Целью работы является исследование процедуры проверки значимости коэффициентов сглаживающего полинома на основе критериев проверки статистических гипотез с целью формирования вектора коэффициентов сглаживающего полинома. Исследовались разработанные методы нелинейного адаптивного сглаживания с оптимизацией степени сглаживающего полинома с оптимизацией структуры сглаживающего полинома. Исследование проводилось путём имитационного моделирования значения вторичных координат, которые по формулам простых методов, пересчитывались в первичные координаты с учетом местоположения и типа измерительных средств. Затем к полученным значениям первичных координат прибавлялись значения ошибок измерений, распределенных по нормальному закону. Полученные таким образом первичные данные измерений подвергались нелинейному адаптивному сглаживанию. Формирование вектора коэффициентов сглаживающего полинома осуществлялось на основе критериев проверки статистических гипотез в следующей последовательности: формирование по данным измерений соответствующей статистики; сравнение этой статистики с пороговым уровнем, зависящим от доверительной вероятности и числа степеней свободы; принятие решения о включении данного компонента в состав полинома. Формирование вектора коэффициентов сглаживающего