Генеративный причинно-следственный подход к анализу данных ИМК

Владимиров Эдуард
vladimirov.ea@phystech.edu

Стрижов Вадим
strijov@phystech.edu

21 декабря 2024 г.

Аннотация

В данной работе предлагается генеративный причинно-следственный подход к анализу взаимодействия биосигналов, в частности электроэнцефалограммы (ЭЭГ) и инерциальных данных. Главная цель — определить и количественно оценить причинно-следственные связи между мозговой активностью и движениями, опираясь на три взаимодополняющие методологии: геометрическое представление сигналов на римановых многообразиях, физически-информированные нейронные сети (PINN) для реалистичного моделирования кинематики, и меры информационной теории и алгоритмы причинного анализа. Объединение этих подходов даёт возможность не только выявлять временные и нелинейные зависимости в системе «мозг–тело», но и строить синтетические выборки для валидации гипотез. Предложенный фреймворк повышает точность и интерпретируемость результатов. Его работа проиллюстрирована на датасете игры с теннисом.

Ключевые слова: ЭЭГ · ИИМ · сходящееся перекрёстное отображение · взаимная информация · причинно-следственный вывод

1 Введение

В последние годы наблюдается стремительный рост интереса к исследованию взаимодействия между мозговой активностью и двигательными действиями человека. С этой целью широко применяются методы анализа электроэнцефалографических (ЭЭГ) сигналов и инерциальных данных (ИИМ, инерциальный измерительный модуль), поскольку они позволяют регистрировать, соответственно, нейронную активность и механику движений в реальном времени. Выявление причинно-следственных связей между ЭЭГ и ИИМ играет важнейшую роль в области нейробиологии, медицины, спортивной науки и технологий Brain-Computer Interface (BCI). Традиционные подходы к анализу таких данных — например, методы корреляции или линейной регрессии — зачастую оказываются недостаточными для точного описания сложных нелинейных взаимодействий в системе «мозг–тело».

Чтобы преодолеть указанные ограничения, в настоящей работе рассматривается **генеративный причинно-следственный подход** (Generative Causal Inference), сочетающий несколько концептуальных направлений:

- **Моделирование сигналов на многообразиях.** ЭЭГ-данные представляют собой высокоразмерный временной ряд, источник которого — множество независимых (или частично независимых) нейронных процессов. Движения, в свою очередь, могут быть представлены на пространстве кинематических параметров, где измерения акселерометра и гироскопа описывают динамику конечностей. Использование геометрических методов и концепции римановой геометрии (Riemannian geometry) позволяет корректно учитывать внутреннюю структуру и ковариации в данных, избегая упрощающих линейных допущений.

- **Физически-информированные нейронные сети (PINN).** Для описания динамики движений и сигналов ИИМ всё чаще применяются Physics-Informed Neural Networks. В отличие от стандартного глубокого обучения, PINN встраивает уравнения классической механики (или другой физической модели) прямо в функцию потерь. Это даёт возможность строить более реалистичные (и интерпретируемые) модели движения, которые затем могут использоваться для проверки гипотез о причинно-следственных связях с мозговой активностью.

- **Причинно-следственный анализ на базе информационных мер.** Использование мер информационной теории, таких как взаимная информация (Mutual Information, MI) и дивергенция Кульбака–Лейблера (Kullback–Leibler divergence, KL), даёт статистически обоснованные способы выявлять и количественно оценивать степень зависимости между сигнальными распределениями. В частности, инструменты вроде Convergent Cross Mapping (CCM) или Probabilistic CCM позволяют проверить, какие каналы ЭЭГ «вызывают» изменения ИИМ, или что связь обусловлена иными факторами.

Объединение подходов моделирования с методами причинно-следственного анализа лежит в основе предлагаемого здесь **генеративного причинно-следственного вывода**. Это позволяет не только установить факт корреляции или временной зависимости между сигналами, но и даёт возможность:

- оценивать устойчивость данных выводов при варьировании параметров модели;
- генерировать синтетические выборки, отражающие ключевые особенности реальных записей ЭЭГ и ИИМ;
- учитывать сложность биологических систем, проявляющуюся в нелинейных и многокомпонентных связях, которые не могут быть адекватно описаны классическими линейными моделями.

Таким образом, целью данной работы является разработка единого методологического фреймворка, который объединяет:

1. геометрические и физические представления о природе исследуемых сигналов,
2. продвинутые методы генерации и анализа данных,
3. алгоритмы выявления причинно-следственных связей.

В дальнейшем мы продемонстрируем, как такой подход может применяться на практике при анализе когнитивных и моторных данных, а также обсудим преимущества и текущие ограничения.

2 Теоретическая часть

2.1 Общие обозначения

Рассмотрим два временных ряда:

$$\mathbf{X}(t) \in \mathbb{R}^{N \times d_X}, \quad \mathbf{Y}(t) \in \mathbb{R}^{N \times d_Y}, \quad t = 1,2,\ldots,T.$$

Пусть $p(\mathbf{X}, \mathbf{Y})$ обозначает совместное распределение этих данных.

Наша цель — проанализировать причинно-следственные связи между $\mathbf{X}$ и $\mathbf{Y}$, то есть понять, существует ли (и насколько выражена) направленная зависимость:

$$\mathbf{X} \longrightarrow \mathbf{Y}, \quad \text{или} \quad \mathbf{Y} \longrightarrow \mathbf{X}.$$

2.2 Независимый анализ компонент (ICA)

При наличии высокоразмерных сигналов $\mathbf{X}(t)$, содержащих множество смешанных источников, часто вводится модель:

$$\mathbf{X}(t) = A\,\mathbf{S}(t),$$

где $\mathbf{S}(t) \in \mathbb{R}^{d_S}$ — вектор независимых компонент, а A — постоянная матрица смешения. Предполагается, что координаты $\mathbf{S}(t)$ статистически независимы:

$$p(\mathbf{S}) \;=\; \prod_{k=1}^{d_S} p\big(S_k\big).$$

Задача независимого анализа компонент сводится к оценке $\widehat{A}^{-1}$ (или непосредственно A) так, чтобы

$$\widehat{\mathbf{S}}(t) \;=\; \widehat{A}^{-1}\,\mathbf{X}(t)$$

максимизировало независимость компонент. Для измерения независимости используют различные функции (энтропию, взаимную информацию и др.). Итоговые независимые компоненты $\widehat{\mathbf{S}}(t)$ далее могут анализироваться вместо $\mathbf{X}(t)$, что часто упрощает выявление причинно-следственных связей, особенно в ЭЭГ-исследованиях.

2.3 Физически-информированные нейронные сети (PINN)

Сигналы $\mathbf{Y}(t)$ могут удовлетворять законам механики, описываемым уравнениями:

$$\frac{d^2\mathbf{r}(t)}{dt^2} = \mathbf{f}\big(\mathbf{r}(t), \dot{\mathbf{r}}(t), \theta\big),$$

где $\mathbf{r}(t) \in \mathbb{R}^m$ — координаты (или углы) исследуемого объекта во времени, а θ — набор физических параметров (масса, длины звеньев и т. п.). При использовании модели PINN мы строим приближающую функцию

$$\mathbf{u}(t,\boldsymbol{\Theta}) \approx \mathbf{r}(t),$$

где $\boldsymbol{\Theta}$ — обучаемые веса нейронной сети. В функцию потерь $\mathcal{L}(\boldsymbol{\Theta})$ помимо ошибки на измеренных данных $\mathbf{Y}(t)$ (которая соответствует ускорению или угловой скорости) добавляют слагаемое, обеспечивающее физическую согласованность:

$$\mathcal{L}(\boldsymbol{\Theta}) = \sum_t \Big\|\ddot{\mathbf{u}}(t,\boldsymbol{\Theta}) - \mathbf{f}\big(\mathbf{u}(t,\boldsymbol{\Theta}),\dot{\mathbf{u}}(t,\boldsymbol{\Theta}),\theta\big)\Big\|^2 + \sum_t \Big\|\ddot{\mathbf{u}}(t,\boldsymbol{\Theta}) - \mathbf{Y}(t)\Big\|^2,$$

где первая сумма отвечает за удовлетворение уравнениям движения, а вторая — за согласие с наблюдениями из IMU.

При введении стохастических факторов (например, случайных начальных условий) получается вероятностная модель

$$p_{\boldsymbol{\Theta}}(\mathbf{Y}) \approx \prod_t p_{\boldsymbol{\Theta}}(\mathbf{Y}(t) \,|\, \mathbf{u}(t,\boldsymbol{\Theta})),$$

позволяющая учесть изменчивость движений при неизменной задаче.

2.4 Метод сходящихся перекрёстных отображений (CCM)

Для проверки гипотезы о том, что $\mathbf{X}(t)$ причинно влияет на $\mathbf{Y}(t)$, используют процедуру Convergent Cross Mapping. Вводится теневое вложение:

$$M_{X,t} = \big(X_t, X_{t-\tau}, \ldots, X_{t-(E-1)\tau}\big),$$

где E — размерность вложения, а τ — временной лаг. Аналогично задаётся $M_{Y,t}$. Если вектор $M_{X,t}$ хорошо *реконструирует* $\mathbf{Y}_t$, то есть

$$\widehat{\mathbf{Y}}_t = \sum_{i=1}^{k} w_i \, \mathbf{Y}_{n_i},$$

где n_i — индексы ближайших соседей $M_{X,t}$ в пространстве M_X, то считается, что есть динамическая причинная связь $\mathbf{X} \to \mathbf{Y}$. Значение

$$\rho_{X\to Y} = \operatorname{corr}\Big(\{\widehat{\mathbf{Y}}_t\},\{\mathbf{Y}_t\}\Big)$$

растёт при увеличении размера библиотеки, если $\mathbf{X} \to \mathbf{Y}$ действительно имеет место.

2.5 Вероятностный метод сходящихся перекрёстных отображений

Вместо единственного предсказания $\widehat{\mathbf{Y}}_t$ можно рассматривать *условное распределение*

$$p_L\big(\mathbf{Y}_t \,\big|\, M_{X,t}\big),$$

оценённое на выборке размера L. По аналогии с CCM, ближайшие соседи используются для построения вероятностной аппроксимации:

$$p_L\big(\mathbf{Y}_t \,\big|\, M_{X,t}\big) \;=\; \sum_{i \in N_L(t)} \alpha_i \, \delta\big(\mathbf{Y} - \mathbf{Y}_{n_i}\big),$$

где $N_L(t)$ — множество индексов соседей, а α_i — веса, зависящие от расстояния в M_X. Тогда для оценки причинности рассматривают, например, величину

$$I_L\big(\mathbf{X} \to \mathbf{Y}\big) \;=\; \sum_t D_{\mathrm{KL}}\Big(p_L\big(\mathbf{Y}_t \,\big|\, M_{X,t}\big) \,\big\|\, p(\mathbf{Y}_t)\Big),$$

или аналогичную меру, связанную с MI (взаимной информацией). При $L \to \infty$ распределение $p_L(\mathbf{Y}_t | M_{X,t})$ должно сходиться к истинному условному $p(\mathbf{Y}_t | M_{X,t})$, давая более строгую вероятностную трактовку динамических причинных связей.

2.6 Ядерное сглаживание

Для набора $\{M_{X,t}\}_{t=1}^T$ можно ввести эмпирическую плотность с помощью ядерного метода. Пусть $K(\cdot)$ — некая ядровая функция (напр. гауссова):

$$\widehat{p}_X(u) \;=\; \frac{1}{T h_X^{m_X}} \sum_{t=1}^{T} K\Big(\frac{\|u - M_{X,t}\|}{h_X}\Big),$$

где $h_X > 0$ — параметр сглаживания, а $m_X = \dim(\mathcal{M}_X)$ — формальная размерность пространства вложения. Аналогично для $M_{Y,t}$ определяем

$$\widehat{p}_Y(v) \;=\; \frac{1}{T h_Y^{m_Y}} \sum_{t=1}^{T} K\Big(\frac{\|v - M_{Y,t}\|}{h_Y}\Big).$$

Таким образом, каждому многообразию $\mathcal{M}_X$ и $\mathcal{M}_Y$ соответствует своя оценка плотности, отражающая эмпирическое распределение вложенных точек.

3 Постановка задачи

Пусть $\mathbf{X}(t) = \{X_1(t), X_2(t), \ldots, X_{n_x}(t)\}$ и $\mathbf{Y}(t) = \{Y_1(t), Y_2(t), \ldots, Y_{n_y}(t)\}$ — два набора многомерных временных рядов, наблюдаемых в моменты времени $t = 1, \ldots, T$. Нас интересует выявление причинных связей между $\mathbf{X}(t)$ и $\mathbf{Y}(t)$, т.е. такие пары связей, где переменные из $\mathbf{X}(t)$ причинно влияют на переменные из $\mathbf{Y}(t)$, или наоборот.

Необходимо определить направленные причинные связи: 1. $X_i(t-\tau) \to Y_j(t)$ для $i = 1, \ldots, n_x$, $j = 1, \ldots, n_y$, и лагов $\tau \geqslant 0$, 2. $Y_j(t-\tau) \to X_i(t)$ для $i = 1, \ldots, n_x$, $j = 1, \ldots, n_y$, и лагов $\tau \geqslant 0$.

Предполагаем, что многомерные временные ряды $\mathbf{X}(t)$ и $\mathbf{Y}(t)$ генерируются следующим образом:

$$X_i(t) = f_i(\mathrm{Pa}_{X_i}(t), \varepsilon_{X_i}(t)),$$
$$Y_j(t) = g_j(\mathrm{Pa}_{Y_j}(t), \varepsilon_{Y_j}(t)),$$

где: - $\mathrm{Pa}_{X_i}(t) \subseteq \{Y_1(t-\tau), \ldots, Y_{n_y}(t-\tau)\}$ — множество родителей переменной $X_i(t)$ из $\mathbf{Y}$, - $\mathrm{Pa}_{Y_j}(t) \subseteq \{X_1(t-\tau), \ldots, X_{n_x}(t-\tau)\}$ — множество родителей переменной $Y_j(t)$ из $\mathbf{X}$, - f_i и g_j — детерминированные функции, описывающие зависимость, - $\varepsilon_{X_i}(t)$ и $\varepsilon_{Y_j}(t)$ — шумовые компоненты.

Оптимизационная задача:

$$\min_{G_{XY}, G_{YX}} \mathcal{L}(\mathbf{X}, \mathbf{Y} \mid G_{XY}, G_{YX}) + \lambda_1 \mathcal{R}(G_{XY}, G_{YX}) + \lambda_2 \mathcal{T}(G_{XY}, G_{YX}),$$

где: - G_{XY} — граф зависимостей $X_i \to Y_j$, - G_{YX} — граф зависимостей $Y_j \to X_i$, - $\mathcal{L}(\mathbf{X}, \mathbf{Y} \mid G_{XY}, G_{YX})$ — правдоподобие наблюдаемых данных с учетом графов G_{XY} и G_{YX}, - $\mathcal{R}(G_{XY}, G_{YX})$ — регуляризатор, штрафующий за сложность графов, - $\mathcal{T}(G_{XY}, G_{YX})$ — штраф за избыточную изменчивость графов во времени.

4 Предлагаемый подход

Algorithm 1 Алгоритм вероятностного выявления влияния $\mathbf{X} \to \mathbf{Y}$ на основе ICA и ядерной оценки плотностей

Входные данные:

- Набор исходных наблюдений $\{\mathbf{X}_{\mathrm{raw}}(t)\}_{t=1}^{T} \subset \mathbb{R}^{d_X}$ (ЭЭГ).
- Набор исходных наблюдений $\{\mathbf{Y}_{\mathrm{raw}}(t)\}_{t=1}^{T} \subset \mathbb{R}^{d_Y}$ (ИИМ).
- Параметры: число независимых компонент r для ICA, размерность вложения E и временной лаг τ, ядровые ширины сглаживания h_X, h_Y.

Выходные данные:

- Временной ряд (или процесс) $\{\gamma(t)\}_{t=1}^{T}$, где

$$\gamma(t) = MI(M_{X,t}, M_{Y,t}) \;=\; \mathbb{E}_{M_{X,t}} D_{\mathrm{KL}}\Big(\widehat{p}\big(M_{Y,t} \,\big|\, M_{X,t}\big) \;\|\; \widehat{p}_Y\big(M_{Y,t}\big)\Big),$$

измеряющий силу влияния $\mathbf{X} \to \mathbf{Y}$ в момент t.

1. **Независимый анализ компонент (ICA).**

   $\mathbf{X}_{\mathrm{raw}}(t) \;=\; A\,\mathbf{S}(t), \quad \mathbf{S}(t) \in \mathbb{R}^{r}, \quad A \in \mathbb{R}^{d_X \times r}.$

   Вычислить оценки независимых компонент:

$$\widehat{\mathbf{S}}(t) \;=\; \widehat{A}^{-1}\,\mathbf{X}_{\mathrm{raw}}(t).$$

2. **Построение временных эмбеддингов для X и Y.**

   Для каждого t сформировать:

$$M_{X,t} \;=\; \big(\widehat{\mathbf{S}}(t), \widehat{\mathbf{S}}(t-\tau), \ldots, \widehat{\mathbf{S}}(t-(E-1)\tau)\big) \;\in\; \mathbb{R}^{rE},$$

$$M_{Y,t} \;=\; \mathbf{Y}_{\mathrm{raw}}(t) \;\in\; \mathbb{R}^{d_Y E}.$$

3. **Ядерная оценка (KDE) для маргинальных плотностей на эмбеддингах.**

   Определить оценки плотностей:

$$\widehat{p}_X(u) \;=\; \frac{1}{T\,h_X^{m_X}} \;\sum_{t=1}^{T} K\Big(\tfrac{\|u-M_{X,t}\|}{h_X}\Big), \quad u \in \mathbb{R}^{m_X},$$

$$\widehat{p}_Y(v) \;=\; \frac{1}{T\,h_Y^{m_Y}} \;\sum_{t=1}^{T} K\Big(\tfrac{\|v-M_{Y,t}\|}{h_Y}\Big), \quad v \in \mathbb{R}^{m_Y},$$

   где $m_X = rE$, $m_Y = d_Y$, а $K(\cdot)$ — ядровая функция.

4. **Агрегация результатов и формирование временного ряда.**

   Сформировать последовательность

$$\gamma(t) \;=\; \mathbb{E}_{M_{X,t}} D_{\mathrm{KL}}\Big(\widehat{p}\big(M_{Y,t} \,\big|\, M_{X,t}\big) \;\|\; \widehat{p}_Y\big(M_{Y,t}\big)\Big), \quad t = 1, \ldots, T.$$