

Множественные связи каналов ЭЭГ и их анализ

Р.И. Ивановский, М.А. Новожилов

Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

Аннотация. Впервые рассмотрены алгоритмы обработки данных ЭЭГ на основе анализа множественных связей. В предложенном подходе последовательно формируются матрицы корреляционных отношений, выделяются доминантные каналы и связанные с ними группы каналов. Рекомендованы параметры для оценки множественных связей внутри групп. Описаны области применения подхода, возможности его использования для выявления информативных корковых зон для отдельных видов психоневрологических заболеваний, при создании новых систем параметрической диагностики этих заболеваний.

Ключевые слова: *корреляционный анализ, регрессионный анализ, множественные связи, обработка ЭЭГ*

ВВЕДЕНИЕ

Развитие подходов к обработке данных электроэнцефалограмм (ЭЭГ) продолжает оставаться весьма актуальным. Это направление во многом определяет возможности модернизации и создания новых образцов медицинской техники, разработки перспективных технологий диагностики, внедрения новых информационных технологий в клиническую практику. Важное значение при обработке имеет анализ межканальных связей в данных ЭЭГ, а, следовательно, связей между отдельными функциональными зонами коры головного мозга (КГМ).

подавляющее большинство отечественных и зарубежных работ в этой области посвящено поиску информативных параметров для анализа парных связей, т. е. взаимодействия двух каналов ЭЭГ между собой. При этом продолжают превалировать работы по анализу связей в рамках линейных моделей. В подобных работах предлагаются, в качестве информативных, различные параметры линейной группы. Среди них преобладающее распространение имеет коэффициент корреляции r Пирсона.

Известно, что область практического применения этого параметра существенно ограничена, однако r служит основой многих прикладных работ по использованию данных ЭЭГ и магнитоэнцефалограмм (МЭГ). Так, в статье [1], с помощью r анализируются межканальные связи по данным ЭЭГ и МЭГ, делается попытка подвести некую базу в обоснование необходимости применения параметра r . Однако авторы [1] используют для этого обоснования явно недостаточные аргументы; они не учитывают, что r служит объективной оценкой уровня связи только для линейных моделей, а простота алгоритма вычисления r не служит гарантией точности получаемых оценок. Ниже будет показано, что параметр r обладает еще одним негативным свойством — низкой чувствительностью к изменчивости уровней связи каналов, а, следовательно, и к изменениям состояния пациентов.

Результаты [1] используются другими исследователями, работающими в рамках линейного подхода. Так, например, на базе этого подхода в работе [2] изучается влияние расстояний между электродами ЭЭГ на качество получаемых данных; в [3]

изучаются различия состояний пациентов в процессе перехода ко сну; в [4] проводится анализ влияния слуховых галлюцинаций на КГМ.

Несмотря на преимущественное распространение коэффициентов корреляции r при анализе данных ЭЭГ, что отмечается в [5], в рамках линейного подхода делаются также попытки применить и другие параметры.

Так, в [6] предложен параметр, названный взаимной информацией $MI(X,Y)$, получение которого требует вычисления объема двумерной плотности распределения вероятностей случайных величин (СВ) X и Y , а также одномерных плотностей распределения этих СВ. Хорошо известно, что оценку плотностей распределений по экспериментальным данным можно осуществить лишь приближенно, на основе соответствующих гистограмм, неизбежно искажающих свойства исходных реализаций. Это всегда связано с появлением дополнительных методических ошибок оценки.

В целом, $MI(X,Y)$ служит лишь еще одним параметром приближенной парной оценки линейной связи, причем свойства $MI(X,Y)$ в [6] не определены. Поэтому утверждение в [7, 8], что параметр $MI(X,Y)$ способен расширить область применения параметра r на нелинейный случай, представляется принципиально неверным. Параметры r и $MI(X,Y)$ относятся к параметрам линейной группы и по этой причине принципиально не имеют возможности расширить области применения друг друга на случаи нелинейных моделей связи.

В [9] применительно к анализу межканальных связей ЭЭГ предложены другие параметры линейной группы — коэффициенты ранговой корреляции Спирмена и Кендалла. Они относятся к непараметрическим коэффициентам корреляции, которые вычисляются по результатам сортировки выборочных данных. Свойства этих параметров, как и $MI(X,Y)$ и r , не выходят за рамки линейных моделей связи. Поэтому, предложенное в [9] сопоставление оценок непараметрических коэффициентов с параметрами $MI(X,Y)$ и r , представляется бесперспективным. Этот вывод можно отнести и к другим попыткам расширить список параметров линейной группы, поскольку выбор очередного нового параметра этой группы для анализа парных связей способен привести лишь к масштабированию получаемых значений.

Обращает на себя внимание редкость работ по анализу парных связей ЭЭГ на основе нелинейных моделей. В работах [10, 11] лишь отмечается необходимость такого анализа связей в данных ЭЭГ. В работе [12] вводится параметр h^2 , названный нелинейным коэффициентом корреляции (НКК). Этот параметр получил обоснование в [13] и подробно описан в [14]. Коэффициент h^2 представлен в этих работах выражением вида:

$$h^2 = \max_{\tau} \left[1 - \frac{z(x(t), \tau)}{\text{var}(y(t + \tau))} \right]; z(x(t), \tau) = \arg \min_g \left(E [y(t + \tau) - g(x(t))]^2 \right). \quad (*)$$

Здесь: $x(t)$, $y(t)$ — временные ряды; var — дисперсия; τ — сдвиг по времени; E — математическое ожидание; $g(t)$ — регрессионное соотношение, представленное в виде кусочно-линейной аппроксимации. Нетрудно видеть, что параметр h^2 определен в упомянутых работах в форме, которая далека от его практического применения: отсутствует алгоритм вычисления h^2 ; нет свидетельств о качестве работы h^2 в условиях нелинейных связей. Этот вывод усугубляется отсутствием описания свойств НКК, а также неопределенностью выбора τ . Действительно, из (*) следует, что рост $|t|$ сопровождается увеличением несинхронности откликов $y(t + \tau)$ и факторов $x(t)$. В результате искусственно искажаются исходные данные. Это неизбежно ведет к появлению дополнительных ошибок оценки уровня связи. Это, по-видимому, объясняет тот факт, что применение параметра h^2 ограничено исследованиями лишь отдельными видами заболеваний, например, эпилепсией [15, 16].

В отечественных публикациях нелинейное направление рассматривается в работе [17], в которой предложен параметр — корреляционное отношение R , лишенный указанных выше недостатков НКК. В [17] обоснована перспективность R , показана простота его применения и эффективность при любых моделях связи. Ниже будут представлены дополнительные достоинства R , связанные с его чувствительностью.

Хорошо известно, что в многоканальных измерительных системах, к которым принадлежит и аппаратура съема данных ЭЭГ, наряду с парными связями, важную роль играют множественные связи, т. е. связи между тремя и более каналами. Перспективность анализа множественных связей очевидна, поскольку такой анализ выявляет связанность ряда зон КГМ между собой, позволяет изучать группирования этих зон по общим признакам. Несмотря на высокую значимость подобного анализа в исследованиях мозга, проблема анализа множественных связей по данным ЭЭГ пока не нашла своего достойного отражения в отечественной и зарубежной литературе. Работ подобного направления применительно к ЭЭГ практически нет.

Среди зарубежных работ отметим лишь статью [12], упомянутую выше. Строго говоря, эта работа не относится к направлению анализа множественных связей в ЭЭГ. Действительно, в ней делается неудачная попытка получить обобщенную оценку связей всех каналов ЭЭГ путем прямого усреднения локальных параметров линейной группы. Очевидно, что подобное усреднение способно привести к абсурдным результатам типа нулевого среднего для корреляций разных знаков. Авторы [12] не учитывают, что решение подобной задачи возможно лишь на основе методов анализа множественных связей.

В отечественных работах необходимость анализа множественных межканальных связей отмечена в упомянутой выше статье [17]

. В ней отмечается появление доминирующих каналов, образование групп вокруг них, указана методическая основа анализа множественных связей.

Завершая краткий анализ современных тенденций в области изучения межканальных связей в данных ЭЭГ, отметим следующее:

- Продолжают превалировать работы в части создания подходов к изучению парных связей на основе линейных моделей, несмотря на известные ограничения области практического применения.
- Дальнейшее расширение списка параметров линейной группы представляется бесперспективным в силу их одинаковой ограниченности. Выбор очередного нового параметра этой группы может привести лишь к масштабированию получаемых значений.
- Предложенный нелинейный коэффициент корреляции h^2 , ввиду его неопределенных свойств, отсутствия рекомендаций к практическому использованию и высокой теоретической сложности, не может рассматриваться в качестве основного параметра при анализе нелинейных связей.
- Наряду с анализом парных связей, большой интерес в общей проблеме изучения мозга имеет взаимодействие нескольких корковых зон, их объединение в группы, что может быть вызвано многими причинами. Выявление и изучение подобных образований требует привлечения специальных методов анализа множественных связей. Несмотря на важность и перспективность подобных задач, публикации по этим вопросам практически отсутствуют.

Дальнейший материал раскрывает сущность анализа множественных связей и содержит описание предлагаемого подхода на этой основе применительно к проблеме изучения межканальных связей ЭЭГ.

ИСХОДНЫЕ ПРЕДПОСЫЛКИ

Пусть данные ЭЭГ получены по m однотипным каналам на промежутке времени t_D , в течение которого фиксируются данные каждого канала последовательностями по n значений в каждой. В результате массив данных ЭЭГ принимает вид $(n \times m)$ -матрицы \mathbf{F} . Значения n и m в конкретной задаче обработки определяются выбранным интервалом времени и числом изучаемых каналов.

Представление исходных данных ЭЭГ в виде матрицы \mathbf{F} универсально. Оно может быть использовано при решении всей совокупности задач обработки: при анализе состояний пациентов в статике и динамике, при любых спектральных преобразованиях данных, произвольных интервалах времени и проч..

Последующее изложение будет опираться на структуру матрицы \mathbf{F} , что подчеркивает общность предлагаемого ниже подхода.

Как отмечалось в [17], анализ парных связей каналов ЭЭГ проводится с использованием различных уравнений (моделей), например, вида:

$$y = a + b \cdot x, \quad (1)$$

$$y = a + b \cdot x + c \cdot x^2, \quad (2)$$

где y и x — отклик и фактор соответственно; a , b и c — коэффициенты, определяемые в задачах регрессии. Модель (1) определяет линейный вариант связи y и x , модель (2) — нелинейный.

Для решения многих задач обработки данных ЭЭГ, в [17] предложено применять корреляционные отношения R и $(m \times m)$ -матрицы \mathbf{R} с элементами R_{ij} ($i, j = 1, \dots, m$). Здесь и далее принято следующее обозначение элементов R_{ij} матриц \mathbf{R} и им подобных: индекс строки (i) элемента R_{ij} определяет номер канала-фактора, а индекс столбца (j) — номер канала-отклика.

В [17] подробно описаны свойства R . Приведем здесь те из них, которые необходимы для последующего изложения. Так, для ненулевых R_{ij} имеем:

- При $R_{ij} = R_{ji}$ связь между каналами линейная вида (1).
- При $R_{ij} \neq R_{ji}$ связь между каналами нелинейная вида (2) и сложнее.
- Выбор большего значения из (R_{ij}, R_{ji}) определяет лучшую модель связи двух каналов. Так, при $R_{ij} > R_{ji}$ лучшая модель реализуется при i -ом факторе и j -ом отклике.

Важно отметить, что выбор R в качестве основного параметра при анализе парных связей объясняется не только способностью R объективно характеризовать связи при любых моделях и определять их лучшие варианты. Параметр R обладает, кроме того, и большей точностью, а, значит, и более высокой чувствительностью к изменениям свойств данных по сравнению с коэффициентами корреляции r . Способность R более заметно реагировать на изменения свойств данных является практически важной. Она позволяет не только более точно оценивать текущее состояние пациентов, но и быстрее отражать изменения состояний по сравнению с r .

Подтвердим это свойство R следующим типовым примером. В качестве исходных рассмотрим данные каналов 8 (фактор) и 10 (отклик) реальной 16-канальной ЭЭГ на интервале времени 20 с. Исходные параметры имели значения: $R = 0.531$, $r = 0.403$, дисперсия фактора $s = 28.451$. С помощью датчика случайных чисел изменим дисперсию фактора, например, до $s = 29.259$. Новые значения R и r составят: $R = 0.518$, $r = 0.399$. Относительные изменения R и r составили: $\delta R = 0.025$, $\delta r = 0.01$, $\delta s = 0.028$. Таким образом, функции чувствительности параметров R и r к данному возмущению равны $\delta R / \delta s = 0.886$ и $\delta r / \delta s = 0.349$ соответственно. Отсюда следует, что параметр R обладает значительно большей (в данном примере — в 2.5 раза) чувствительностью, чем r . Интерактивный ресурс с решением подобного примера размещен на портале [18]

в рубрике «Математика. Математика в приложениях» под номером 6 и называется «Анализ связей каналов ЭЭГ. Сравнение чувствительности КО и КК».

Наряду с парными связями, в многоканальных системах могут иметь место связи внутри локальных групп каналов. Выделение таких групп и их анализ по данным ЭЭГ представляет несомненный интерес, поскольку это позволит изучить процессы группирования корковых зон, уточнить общую информационную структуру коры головного мозга. Возможность появления подобных групп при анализе данных ЭЭГ впервые отмечена в работе [17]. В ней показано, что выделение в матрицах \mathbf{R} превалирующих элементов из пар (R_{ij}, R_{ji}) приводит к образованию столбцов с несколькими (не менее двух) элементами. Номер каждого из таких столбцов определяет номер доминирующего канала (отклика), а индексы i элементов этого столбца — номера каналов-факторов, влияющих на доминирующий канал. В результате образуются группы каналов, которые назовем D -группами. Оставляя в структуре матрицы \mathbf{R} только превалирующие элементы из пар (R_{ij}, R_{ji}) , получаем матрицу \mathbf{R}_0 , каждый столбец которой характеризует определенную D -группу. Общее число D -групп в матрице \mathbf{R} не превышает m .

Таким образом, каждую D -группу составляют s каналов ($s \geq 3$), в том числе доминантный (основной) канал-отклик с данными y и связанные с ним каналы-факторы с данными $(x_1, x_2, \dots, x_{s-1})$. Здесь символы y и x — значения соответствующих случайных величин (СВ).

На рис. 1 приведена матрица \mathbf{R} и полученная из нее матрица \mathbf{R}_0 D -групп. Исходными при этом служили данные реальной ЭЭГ, матрица \mathbf{F} для которых имеет параметры: $m = 16$, $n = 11486$, $t_D = 45.22$ с. Расчеты проводились при $t_D = 45$ с. Матрица \mathbf{R} (см. рис. 1) сформирована в рамках модели (2) для совокупности каналов с 6 по 11. Единицы в диагоналях \mathbf{R} и \mathbf{R}_0 заменены номерами каналов для удобства работы с элементами этих матриц.

6	0.262	0.364	0.152	0.506	0.468
0.261	7	0.374	0.548	0.441	0.578
0.438	0.413	8	0.258	0.597	0.398
0.156	0.561	0.19	9	0.355	0.472
0.508	0.433	0.464	0.349	10	0.511
0.47	0.58	0.35	0.438	0.511	11

6	0	0	0	0	0
0	7	0	0	0.441	0
0.438	0.413	8	0.258	0.597	0.398
0.156	0.561	0	9	0.355	0.472
0	0	0	0	10	0
0	0	0	0	0	11

Рис. 1. Матрицы \mathbf{R} и \mathbf{R}_0 для группы из шести каналов

Из приведенной матрицы \mathbf{R}_0 следует, что в данной совокупности каналов имеются четыре D -группы с основными каналами 6, 7, 10 и 11. Десятый канал имеет три канала-фактора, каналы 6, 7, 11 — по два.

Интересно отметить (см. рис. 1), что все D -группы матрицы \mathbf{R}_0 включают в состав факторов каналы 8 и 9, которые в данном примере являются своеобразными информационными генераторами. Подобные совокупности каналов-факторов, влияющих на несколько каналов-откликов, назовем G -группой. Сделанное замечание может иметь прикладное значение и заслуживает отдельного рассмотрения.

Предварительный анализ данных многих ЭЭГ показал высокую распространенность D -групп, что говорит о наличии еще одной структуры межканальных связей. Определим подход, необходимый для изучения этой структуры и составляющих ее D -групп.

Ниже излагается основа такого подхода, который использует элементы теории множественной регрессии.

АНАЛИЗ D -ГРУПП

Случаи, когда на отклик влияют несколько факторов, относятся к специальным разделам математической статистики, посвященным анализу множественных связей. Свойства множественных связей в составе D -групп описывают несколько параметров, значения которых могут рассматриваться в качестве индивидуальных показателей состояния пациента. В состав этих показателей следует включить и сам факт наличия определенной D -группы, объединяемые ею зоны, а также число и номера каналов-факторов.

Интегральный уровень связей в каждой D -группе между основным ее каналом (откликом) и всей совокупностью его каналов-факторов характеризует сводный коэффициент корреляции R_s . Согласно определению, [19], параметр R_s есть парный коэффициент корреляции отклика и его оценки. Значение R_s может быть получено двумя путями.

В первом [20] предложен алгоритм получения R_s без явного решения экстремальной задачи регрессии. Эта процедура использует данные отклика и факторов во вспомогательных векторах и матрицах. Оценка R_s получается в результате последовательности операций с минорами вспомогательных матриц. Такой алгоритм сложен в практическом применении, требует формального подхода при реализации.

Другой, значительно более простой и удобный, алгоритм оценки R_s [19] прямо использует определение этого параметра, приведенное выше. Для получения R_s предварительно решается задача множественной регрессии, формируются векторы откликов \mathbf{Y} и его оценок \mathbf{Y}_0 , вычисляется коэффициент корреляции этих векторов.

Из определения параметра R_s и алгоритмов его оценки следует, что этот параметр является пострегрессионным, т. е. для его оценки требуется решение задачи регрессии (в данном случае — множественной). Это сближает свойства R_s со свойствами корреляционного отношения R в части возможности применения при любых моделях связей и использования в качестве критерия при выборе лучших вариантов моделей. Полный анализ близости свойств параметров R_s и R выходит за рамки настоящего материала. Отметим лишь, что эти параметры могут быть вычислены по одним и тем же формулам, так в задачах с одним фактором ($s = 2$) значение R_s равно R .

Поскольку оценка R_s тесно связана с решением задачи множественной регрессии (ЗМР), рассмотрим ее некоторые особенности. Как и в однофакторных задачах, в ЗМР необходимо задать регрессионное соотношение, используя в нем один из вариантов (линейный или нелинейный) предполагаемой модели связи. Для определенности рассмотрим ЗМР для D -группы с двумя факторами x и z . Тогда, в качестве регрессионных, запишем уравнения, соответственно для линейной:

$$y = a + b \cdot x + c \cdot z; \quad (3)$$

и нелинейной модели связи:

$$y = a + b \cdot x + c \cdot x^2 + d \cdot z + g \cdot z^2. \quad (4)$$

где a, b, c, d, g — коэффициенты, подлежащие определению в ЗМР.

Используя выражения типа (3) и (4) в качестве регрессионных, получаем линейную ЗМР, поскольку отклик y линейно зависит от искоемых коэффициентов (a, b, c и др.). Решение любой линейной ЗМР может быть достигнуто методом наименьших квадратов (МНК). Приведем основные соотношения алгоритма МНК, пользуясь его формой из работы [19].

Данные ЭЭГ рассматриваемой D -группы находятся в $s = 3$ соответствующих столбцах матрицы \mathbf{F} . Один из столбцов содержит данные отклика y_i , другие — факторов x_i, z_i ($i = 1, 2, \dots, n$). Тогда, для каждой i -ой совокупности отклик-факторы, согласно, например, модели (4), имеем:

$$y_i = a + b \cdot x_i + c \cdot (x_i)^2 + d \cdot z_i + g \cdot (z_i)^2 = [1 \ x_i \ (x_i)^2 \ z_i \ (z_i)^2] \cdot \beta = \mathbf{a}_i \cdot \beta. \quad (5)$$

Здесь: \mathbf{a}_i — i -ая строка вспомогательной матрицы \mathbf{A} ; β — вектор искомых коэффициентов, в данном случае $\beta = [a \ b \ c \ d \ g]^T$; T — символ операции транспонирования. Для всех n уравнений (5) получаем:

$$\mathbf{Y} = \mathbf{A} \cdot \beta, \quad (6)$$

где \mathbf{Y} — $(n \times 1)$ -вектор откликов с элементами y_i ; вспомогательная матрица \mathbf{A} состоит из строк \mathbf{a}_i и, для модели (4), имеет размерность $(n \times 5)$.

В процессе решения задачи регрессии последовательно определяются: вектор оценок искомых коэффициентов β_0 , вектор \mathbf{Y}_0 оценок откликов, остаточная сумма квадратов ε и остаточная дисперсия d :

$$\beta_0 = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{Y}; \mathbf{Y}_0 = \mathbf{A} \cdot \beta_0; \mathbf{e} = \mathbf{Y} - \mathbf{Y}_0; \varepsilon = \mathbf{e}^T \cdot \mathbf{e}; d = \varepsilon / n. \quad (7)$$

Качество решения ЗМР, как и других задач регрессии, зависит от вида регрессионных соотношений. Оценить это качество позволяют результирующие значения параметров ε и d , которые минимизируются алгоритмом МНК. Чем меньше ε и d , тем теснее будет связь векторов \mathbf{Y} , \mathbf{Y}_0 и тем выше будет значение параметра R_s . Поэтому сводный коэффициент корреляции R_s может служить критерием согласованности получаемых оценок с исходными данными. Чем выше значение R_s , тем лучше выбранная модель связи отражает свойства исходных данных. Завершая общую характеристику сводного коэффициента корреляции, отметим, что значения $R_s \geq 0$, поскольку векторы \mathbf{Y} и \mathbf{Y}_0 всегда согласованы между собой.

В табл.1 приведены значения R_s , полученных для D -групп ($s = 3$) из матрицы \mathbf{R}_0 (см. рис.1) при использовании моделей первого (3) и второго (4) порядков. Из полученных значений R_s следует, что и для множественных связей нелинейная модель полнее отражает свойства исходных данных ЭЭГ.

Таблица 1

Основной канал	Модель (3)	Модель (4)
6	0.423	0.515
7	0.613	0.63
11	0.515	0.555

Оценка R_s служит обобщенной характеристикой конкретной D -группы, позволяет характеризовать свойства множественных связей в D -группе одним параметром. Поэтому параметр R_s удобно использовать для контроля текущего состояния пациента, оценки динамики изменения состояний и решения многих других прикладных задач.

Интерактивный ресурс с получением значений R_s размещен на портале [18] в рубрике «Математика. Математика в приложениях» под номером 5 и называется «Анализ связей каналов ЭЭГ. D-группы».

Наряду со сводным коэффициентом корреляции, при анализе D -групп могут быть использованы и другие параметры множественных связей. Так, множественная регрессия позволяет рассмотреть связь двух каналов, каждый из которых подвержен влиянию одной и той же совокупности других каналов. Этот случай относится к, так называемой, частной корреляции, основным параметром которой служит частный коэффициент корреляции R_c .

Рассмотрим D -группу, в состав которой входят основной канал-отклик с данными y и связанные с ним каналы-факторы с данными x_1, x_2, \dots, x_{s-1} (здесь y и x — значения соответствующих случайных величин).

Пусть, например, анализируется связь y и x_1 , причем каждая из этих СВ находится под влиянием остальных факторов, т. е. $(x_2, x_3, \dots, x_{s-1}) = Q$. При этом возникают две

ЗМР с откликами y и x_1 , и факторами Q в каждой. В результате решения этих двух задач будут определены оценки откликов y и x_1 , а также ошибки e_0 и e_1 этих оценок. Тогда частным коэффициентом корреляции для y и x_1 называется парный коэффициент корреляции e_0 и e_1 .

Аналогично R_s , оценка частного коэффициента корреляции R_c может быть получена двумя путями. Первый из них предлагает использовать формальный алгоритм [20], который основан на формировании вспомогательных векторов и матриц из данных откликов и факторов. Оценка R_c является результатом сложных операций с минорами вспомогательных матриц, причем при смене вариантов анализа индексация миноров комбинаторно меняется. Применение такого алгоритма на практике затруднительно вследствие высокой степени формализма.

В современных условиях развития компьютерных технологий оценка R_c может быть получена значительно проще, путем прямого вычисления ошибок оценки (в нашем примере — e_0 и e_1) и получения для них коэффициента корреляции [19]. Основу такого алгоритма составляет решение двух ЗМР. Так, для рассматриваемого примера, в первой задаче участвуют отклик y и факторы Q ; во второй задаче откликом служит x_1 , факторами — те же Q . В результате для этого варианта получаем оценку частного коэффициента корреляции y и x_1 . Переход к другим вариантам пар внутри той же группы потребует лишь замены исходных данных в ЗМР.

К достоинствам второго алгоритма, кроме простоты и удобства, следует отнести также возможность выбора лучших вариантов моделей для каждой из двух решаемых ЗМР. Лучшей модели, как отмечалось выше, соответствуют наименьшие значения остаточной суммы квадратов ϵ или остаточной дисперсии d . Критериями качества решения могут служить также сводные коэффициента корреляции в каждой из ЗМР. Значения частных коэффициентов корреляции R_c , в отличие от R_s , могут быть и отрицательными, поскольку R_c отражает корреляцию двух разнородных СВ.

В табл.2 приведены значения частных коэффициентов корреляции R_c , полученных для двухфакторных D -групп (см. рис.1) при использовании в ЗМР моделей первого порядка (3). Факторами служили данные каналов 8 и 9. Второй столбец табл. 2 характеризует корреляцию основных каналов с каналом-фактором 9; третий столбец — с каналом-фактором 8.

Таблица 2.

Основной канал	Канал 9	Канал 8
6	-0.235	-0.4
7	-0.524	0.336
11	-0.408	0.302

В табл. 3 для тех же каналов приведены коэффициенты корреляции r , сопоставление которых с данными табл. 2 показывает роль дополнительных факторов в усилении (для канала 6) или ослаблении (для каналов 7 и 11) простой парной связи.

Таблица 3.

Основной канал	Канал 9	Канал 8
6	-0.151	-0.363
7	-0.544	0.374
11	-0.438	0.345

Сопоставляя рассмотренные параметры R_s и R_c , обратим внимание на большее удобство R_s в практическом применении. Действительно, значение R_s дает общую оценку множественных связей D -группы. Единственный параметр D -группы проще контролировать, чем значения множества R_c , число которых может достигать $(m - 1)$. В

то же время отметим, что с помощью R_c можно анализировать детали множественных связей D -групп.

Эти соображения позволяют рекомендовать параметры R_s , в качестве основных, при анализе множественных связей в данных ЭЭГ, оценки состояний пациентов, получении пороговых значений, отделяющих одно состояние от другого, а также при решении многих других прикладных задач.

ВОЗМОЖНЫЕ ПРИЛОЖЕНИЯ

Описанный подход, основанный на анализе множественных связей, использовании D -групп и сводного коэффициента корреляции R_s , может послужить базой для создания ряда технологий решения таких прикладных задач, как, например:

- классификации вида заболевания (T_1);
- оценки текущего состояния пациента (T_2);
- оценки динамики изменения состояний, построения моделей динамики и прогнозирования состояний на их основе (T_3).

Подробное описание здесь этих технологий потребует значительного увеличения объема данного материала. Поэтому остановимся здесь лишь на сущности их основных этапов.

Важнейшими этапами T_1 являются: подготовка и отбор контрольных групп (например, Гр1 — здоровых, Гр2 — с определенным видом заболевания); получение данных ЭЭГ; получение матриц \mathbf{R}_0 и значений R_s для каждого столбца \mathbf{R}_0 каждого из участников Гр1 и Гр2; получение точечных и интервальных оценок R_s для доминантных каналов в Гр1 и Гр2; выявление информативных корковых зон, т. е. зон, для которых отличия значений R_s в Гр2 от значений R_s в Гр1 проявились особенно резко. Результатами выполнения этих этапов служат конкретные зоны, а также диапазоны значений их сводных коэффициентов для определенного вида заболевания. Данная методика предполагает необходимость тщательного отбора участников Гр1, Гр2 и соблюдения одинаковых условий съема данных ЭЭГ для обеспечения их однородности. Ниже представлены результаты предварительной апробации T_1 .

В рамках T_2 рассматривается пациент с известным видом заболевания, информативные зоны для которого предполагаются известными по результатам T_1 . Для этих зон определяются значения R_s . Результаты используются для оценки текущего состояния и принятых решений по программе лечения.

Сущность T_3 заключается в периодической оценке R_s , построении динамической модели в виде системы дифференциальных уравнений и использования их для прогнозирования дальнейшего развития процессов. Подобные модели могут применяться для решения широкого круга прикладных задач, включая, например, задачу определения требуемых уровней дозирования конкретных препаратов. В целом, подобный подход нуждается в отдельном детальном рассмотрении и изложении.

Среди задач перечисленных типов особое место занимают задачи T_1 , результаты решения которых, ввиду многократного последующего использования, должны обладать необходимой объективностью.

Это, в свою очередь, требует высокого качества исходных данных, достижение которого возможно при выполнении ряда условий:

1. число участников каждой из Гр1 и Гр2 должно быть не менее 40;
2. в целях снижения разброса данных необходимо стремиться к обеспечению однородности состава каждой из контрольных групп;
3. в этих же целях съем данных ЭЭГ должен производиться для каждого из участников контрольных групп в единых условиях с высокой тщательностью.

Последняя позиция нуждается в пояснении.

При проведении серий любых экспериментов важно соблюдать неизменность условий проведения. Это в полной мере относится и к массовому съему данных ЭЭГ. На качество данных влияет множество факторов. Среди них состояние пациента, его напряженность, внешние возмущения; погрешности позиционирования электродов, плотность контактов с кожным покровом и т. д.

Особую сложность представляет, по-видимому, поддержание состояния каждого пациента контрольной группы на одном уровне. Достижению последнего могли бы способствовать дополнительные меры, связанные: с парированием визуальных и слуховых возмущений (повязка на глаза, накладные наушники каждому пациенту перед ЭЭГ); со стабилизацией в процессе съема ЭЭГ внутреннего состояния пациентов на одном уровне (например, считать вслух цифры в темпе, указанном тихими щелчками метронома в наушниках). При обнаружении значительных разбросов, данные так же могут быть подвергнуты заключительной корректировке на основе методов обеспечения однородности выборок [19].

НЕКОТОРЫЕ ПРЕДВАРИТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

В процессе подготовки данного материала был проведен экспресс-анализ информативности описанного подхода. Исходными служили массивы данных¹ 16-канальных ЭЭГ для двух контрольных групп (Гр1 — здоровые люди, Гр2 — с болезнью Паркинсона) численностью 21 чел. каждая. Исходные данные содержали также информацию по принятой схеме соответствия номеров каналов кодам конкретных функциональных зон КГМ.

Согласно технологии T_1 , были последовательно получены матрицы \mathbf{R}_0 и значения сводных коэффициентов R_s для каждого столбца \mathbf{R}_0 каждого из участников Гр1 и Гр2. В качестве точечных оценок были вычислены средние значения $R_{ср}$ по множествам одноименных столбцов каждой из Гр1 и Гр2. Интервальные оценки R_s для каналов в данном случае не определялись ввиду недостаточной численности Гр1 и Гр2.

На рис. 2 приведена матрица \mathbf{M} , объединяющая полученные результаты.

$$\mathbf{M} = \begin{array}{c} \begin{array}{cccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ \hline 0.742 & 0.792 & 0.801 & 0.777 & 0.811 & 0.811 & 0.789 & 0.752 & 0.795 & 0.787 & 0.77 & 0.622 & 0.757 & 0.753 & 0.765 & 0.735 \\ 0.776 & 0.71 & 0.763 & 0.738 & 0.701 & 0.731 & 0.723 & 0.776 & 0.784 & 0.761 & 0.754 & 0.717 & 0.764 & 0.703 & 0.705 & 0.733 \\ 0.046 & -0.104 & -0.047 & -0.05 & -0.136 & -0.099 & -0.084 & 0.032 & -0.014 & -0.033 & -0.021 & 0.153 & 9.247 \cdot 10^{-3} & -0.066 & -0.078 & -2.721 \cdot 10^{-3} \end{array} \end{array}$$

Рис. 2. Результаты анализа

Первая строка матрицы \mathbf{M} содержит номера каналов, вторая (вектор \mathbf{n}) и третья (вектор \mathbf{p}) строки — значения $R_{ср}$ для Гр1 (здоровых) и Гр2 (больных) соответственно. В последней строке матрицы \mathbf{M} приведены относительные отклонения значений элементов вектора \mathbf{p} от значений элементов вектора \mathbf{n} .

Из данных рис. 2 следует, что в рассматриваемом примере наибольшие отличия значений $R_{ср}$ в Гр2 от Гр1 проявились для каналов 2, 5 и 12.

Этим каналам соответствуют [21] функциональные зоны КГМ Fp2, C3 и F8. Согласно описанию функциональных характеристик этих зон [21, 22], они имеют непосредственную связь с симптомами болезни Паркинсона [23, 24], например, в части снижения мышечной активности, нарушений в походке и др..

Оценивая полученные результаты, отметим, что исходные данные в проведенном анализе характеризовались малой численностью контрольных групп. К ним не применялось мер по повышению однородности выборок. Но, даже в этих условиях

¹ Данные любезно предоставлены к.м.н. А.Л.Гореликом, за что авторы выражают А.Л.Горелику искреннюю благодарность

применение предлагаемого подхода позволило выявить локальную группу информативных функциональных зон для рассмотренного вида заболевания. Проведенный анализ можно рассматривать в качестве иллюстративного примера практического применения параметров множественной корреляции и предложенных алгоритмов обработки данных ЭЭГ. Последующая апробация описанного подхода в клинической практике, решение на этой основе прикладных задач позволит уточнить сделанные выводы. Подтверждение высказанных выше предположений послужит основанием для разработки новых технологий параметрической диагностики конкретных психоневрологических заболеваний по локальным группам информативных для них функциональных зон КГМ.

ВЫВОДЫ

В статье впервые, применительно к проблеме обработки и использования данных ЭЭГ, предлагается подход, основанный на теории множественной регрессии. Этот подход служит развитием ранее разработанной авторами технологии обработки данных ЭЭГ с применением матриц \mathbf{R} корреляционных отношений. Показано, что структура матриц \mathbf{R} содержит совокупность доминантных групп (D -групп), анализ которых опирается на теорию множественной регрессии. Предложены параметры множественных связей D -групп и алгоритмы их получения. В качестве основного параметра множественных связей рекомендуется сводный коэффициент корреляции R_s , обладающий рядом потенциально важных достоинств. Предложены основные направления практического применения разработанного подхода. Приведены результаты предварительной апробации предложенной технологии применительно к случаю паркинсонизма. Разработанный подход может служить основой создания новой технологии выявления информативных локальных групп функциональных зон КГМ по отдельным видам психоневрологических заболеваний. Это позволит практически использовать параметрическую диагностику этих заболеваний по их локальным функциональным зонам.

СПИСОК ЛИТЕРАТУРЫ

1. Cover K.S., Stam C.J., van Dijk B.W. Detection of very high correlation in the alpha band between temporal regions of the human brain using MEG. *Neuroimage*. 2004. V. 22. № 4. P. 1432-1437. doi: [10.1016/j.neuroimage.2004.04.016](https://doi.org/10.1016/j.neuroimage.2004.04.016)
2. Ramakrishnan A.G., Satyanarayana J.V. Reconstruction of EEG from limited channel acquisition using estimated signal correlation. *Biomedical Signal Processing and Control*. 2016. V. 27. P. 164-173. doi: [10.1016/j.bspc.2016.02.004](https://doi.org/10.1016/j.bspc.2016.02.004)
3. Kalauzi A., Vuckovic A., Bojić T. EEG alpha phase shifts during transition from wakefulness to drowsiness. *International journal of psychophysiology*. 2012. V. 86. № 3. P. 195-205. doi: [10.1016/j.ijpsycho.2012.04.012](https://doi.org/10.1016/j.ijpsycho.2012.04.012)
4. Angelopoulos E., Koutsoukos E., Maillis A., Papadimitriou G.N., Stefanis C. Cortical interactions during the experience of auditory verbal hallucinations. *The Journal of neuropsychiatry and clinical neurosciences*. 2011. V. 23. № 3. P. 287-293.
5. Vega-Zelaya L., Garnés-Camarena O., Sanz-García A., Ortega G. J., Pastor J. *Clinical Advances in Neurophysiology*. 2016.
6. Ventouras E.C., Margariti A., Chondraki P., Kalatzis I., Economou N.T., Tsekou H., Paparrigopoulos T., Ktonas P. EEG-based investigation of brain connectivity changes in psychotic patients undergoing the primitive expression form of dance therapy: a methodological pilot study. *Cognitive neurodynamics*. 2015. V. 9. № 2. P. 231-248. doi: [10.1007/s11571-014-9319-8](https://doi.org/10.1007/s11571-014-9319-8)

7. Alonso J.F., Mananas M.A., Romero S., Hoyer D., Riba J., Barbanoj M.J. Drug effect on EEG connectivity assessed by linear and nonlinear couplings. *Human brain mapping*. 2010. V. 31. № 3. P. 487-497. doi: [10.1002/hbm.20881](https://doi.org/10.1002/hbm.20881)
8. Cruz A.V., Mallet N., Magill P.J., Brown P., Averbach B.B. Effects of dopamine depletion on information flow between the subthalamic nucleus and external globus pallidus. *Journal of neurophysiology*. 2011. V. 106. № 4. P. 2012-2023. doi: [10.1152/jn.00094.2011](https://doi.org/10.1152/jn.00094.2011)
9. Bonita J.D., Ambolode II L.C.C., Rosenberg B.M., Cellucci C.J., Watanabe T.A.A., Rapp P.E., Albano A.M. Time domain measures of inter-channel EEG correlations: a comparison of linear, nonparametric and nonlinear measures. *Cognitive neurodynamics*. 2014. V. 8. № 1. P. 1-15. doi: [10.1007/s11571-013-9267-8](https://doi.org/10.1007/s11571-013-9267-8)
10. Jeong J., Kim D.J., Chae J.H., Kim S.Y., Ko H.J., Paik I.H. Nonlinear analysis of the EEG of schizophrenics with optimal embedding dimension. *Medical engineering & physics*. 1998. V. 20. № 9. P. 669-676. doi: [10.1016/S1350-4533\(98\)00078-2](https://doi.org/10.1016/S1350-4533(98)00078-2)
11. Meyer-Lindenberg A., Bauer U., Krieger S., Lis S., Vehmeier K., Schüler G., Gallhofer B. The topography of non-linear cortical dynamics at rest, in mental calculation and moving shape perception. *Brain topography*. 1998. V. 10. № 4. P. 291-299. doi: [10.1023/A:1022227108139](https://doi.org/10.1023/A:1022227108139)
12. da Silva F.L., Pijn J.P., Boeijinga P. Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts. *Brain topography*. 1989. V. 2. № 1-2. P. 9-18. doi: [10.1007/BF01128839](https://doi.org/10.1007/BF01128839)
13. Kalitzin S.N., Parra J., Velis D.N., da Silva, F.L. Quantification of unidirectional nonlinear associations between multidimensional signals. *IEEE transactions on biomedical engineering*. 2007. V. 54. № 3. P. 454-461. doi: [10.1109/TBME.2006.888828](https://doi.org/10.1109/TBME.2006.888828)
14. Wendling F., Ansari-Asl K., Bartolomei F., Senhadji L. From EEG signals to brain connectivity: a model-based evaluation of interdependence measures. *Journal of neuroscience methods*. 2009. V. 183. № 1. P. 9-18. doi: [10.1016/j.jneumeth.2009.04.021](https://doi.org/10.1016/j.jneumeth.2009.04.021)
15. Pijn JPM (1990) Quantitative evaluation of EEG signals in epilepsy: nonlinear associations, time delays and nonlinear dynamics. PhD thesis, University of Amsterdam. 1990.
16. Wendling, F., Bartolomei, F., Bellanger, J. J., Chauvel, P. Interpretation of interdependencies in epileptic signals using a macroscopic physiological model of the EEG. *Clinical neurophysiology*. 2001. V. 112. № 7. P. 1201-1218. doi: [10.1016/S1388-2457\(01\)00547-8](https://doi.org/10.1016/S1388-2457(01)00547-8)
17. Ивановский Р. И., Новожилов М. А. Анализ межканальных связей электроэнцефалограмм на основе корреляционных отношений // *Математическая биология и биоинформатика*. 2016. Т. 11. № 2. С. 214-224. doi: [10.17537/2016.11.214](https://doi.org/10.17537/2016.11.214)
18. *MAS.Exponenta.ru*. URL: <http://mas.exponenta.ru/> (accessed 25.04.2017).
19. Ивановский Р.И. *Теория вероятностей и математическая статистика. Основы, прикладные аспекты с примерами и задачами в среде Mathcad*. СПб: БХВ, 2008. 528 с.
20. Крамер Г. *Математические методы статистики*. / Пер. с англ. под ред. А. Н. Колмогорова. М.: Мир, 1975. 648 с. Translation of: Cramer H. *Mathematical methods of statistics*. Uppsala: Almqvist & Wiksells, 1945. 575 p.
21. Шелякин А. М., Преображенская И. Г., Богданов О. В. *Микрополяризационная терапия в детской неврологии*. М.: Медкнига, 2008. 120 с.
22. Триумфов А. В. *Топическая диагностика заболеваний нервной системы*. М: МЕДпресс-информ, 2015. 279 с.

23. Simon R. P., Greenberg D. A., Aminoff M. J. *Clinical neurology*. Lange Medical Books/McGraw-Hill, 2009.
24. Гусев Е. И., Коновалов А. Н., Бурд Г. С. *Неврология и нейрохирургия: учебник*. М.: Медицина, 2000. 374 с.