И.В. Доровских, А.Ю. Щеглов, А.Г. Любин

ВЫЧИСЛИТЕЛЬНАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ ПСИХИЧЕСКИХ НАРУШЕНИЙ ОСТРОГО ПЕРИОДА СОТРЯСЕНИЯ ГОЛОВНОГО МОЗГА

Введение

Исследовательский интерес к легкой черепно-мозговой травме, в частности, к сотрясению головного мозга (СГМ), в настоящее время ощутимо возрастает. Это определяется абсолютным преобладанием СГМ в структуре нейротравматизма, до 90% [1]. Вместе с тем, психические нарушения острого периода СГМ остаются вне поля зрения ввиду незначительной длительности и скоротечности нарушений. Кроме того, современное нейропсихиатрическое исследование сопряжено с большими объемами собираемой и обрабатываемой информации. Очевидно, что создаваемые при этом информационные массивы нуждаются в серьезной вычислительной обработке с целью повышения информативности, доступности, а также для снижения времени на принятие по ним диагностического и экспертного решения. Собранные базы данных могут использоваться для постановки и решения ряда более общих задач, связанных с выведением общих и средних показателей по диагностике психического нарушениястепени восстановления функций мозга после лечения, и, прежде всего, с организацией автоматического поиска аналогов по уже обследованным пациентам, что безусловно является актуальной задачей.

Аппроксимация серийных обследований

При каждом обследовании функционального и психического состояния каждого пациента в стационаре формируется вектор $\bar{x} \in R^n$, состоящий иэ N числовых (вещественных) параметров. На практике $N \geq 200$. За время пребывания в клинике подавляющее большинство пациентов проходит три полных обследования. Таким образом, общее число компонент вектора $\bar{x}$ утраивается. При обычной, неавтоматизированной практике большинство получаемых данных обследований заносятся в карту пациента, хранятся в ней и редко востребуются для анализа. Это легко объясняется труднодоступностью больших объемой рукописной числовой информации.

Ряд проводимых исследований и получаемых параметров носит серийный характер. К такого рода данным можно отнести, например, результаты оценивания пациентами в различных режимах заданных временных интервалов. При обычной практике эта информация формирует ряд значений (координат) вектора $\vec{x}$. По существу, такие массивы данных - это значения функции, заданные на некотором временном отрезке. Обычно область определения такой функции – это отрезок $D = [2, 60]$ (в секундах). Для удобства восприятия предлагается, во-первых, рассматривать относительное отклонение $y_j = (z_j - z_j^T)/z_j^T$ экспериментальных данных z_j от точных значений z_j^T вместо абсолютных значений z_j, $j=1,2,...,n$, получаемых экспериментально. Такой переход позволяет привести к одному масштабу данные, соответствующие малым и большим временным интервалам и сделать их восприятие более сопоставимым. Во-вторых, предлагается проводить аппроксимацию этих экспериментальных данных на области определения фрагментом параболы с выбираемыми оптимальным образом (через минимизацию квадратичной невязки) тремя параметрами, определяющими параболу. Такого вида приближения можно отнести к однородным квадратичным сплайнам. Параметры аппроксимирующей кривой при этом вычисляются аналитически. Так, если обозначить через y_i значения приближенно заданной функции в точках отрезка D и через $f(t) = at^2 + bt + c$ -- выражение для аппроксимирующей функции с параметрами a, b, c, то задача состоит в минимизации в норме пространства $L_2(D)$ или R^n невязки, записываемой, например, в виде

$$J(a,b,c) = \sum_{j=1}^{n} \left| f(t_j) - y_j \right|^2 . \tag{1}$$

При этом из условий существования [2] глобального минимума квадратичной невязки (1) следуют явные формулы для вычисления компонент вектора $\arg\min J(a,b,c)$:

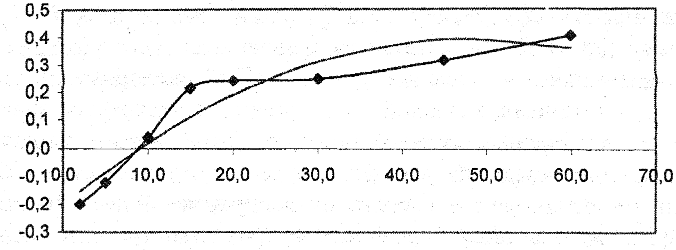
$$a = (e_0d_1d_3 + e_1d_1d_2 + e_2d_0d_2 - e_0d_2^2 - e_1d_0d_3 - e_2d_1^2)\, S^{-1},$$
$$b = (e_0d_2d_3 + e_1d_0d_4 + e_2d_1d_2 - e_0d_1d_4 - e_1d_2^2 - e_2d_0d_3)\, S^{-1},$$
$$c = (e_0d_2d_4 + e_1d_2d_3 + e_2d_1d_3 - e_0d_3^2 - e_1d_1d_4 - e_2d_2^2)\, S^{-1},$$

где $S = 2d_1d_2d_3 + d_0d_2d_4 - d_0d_3^2 - d_1^2d_4 - d_2^3$ при

$$d_i = \sum_{k=1}^{n} t_k^i , \qquad e_j = \sum_{k=1}^{n} t_k^j y_k , \qquad i=0,1,2,3,4, \qquad j=0,1,2.$$

В качестве примера может быть представлена следующая иллюстрация (см. рис. 1).

Рис 1. Относительные отклонения воспроизведения временных интервалов



На графике сплошной линией с ромбами представлена информация, полученная при первом обследовании больного. Эта информация имеет дискретный характер, и на рисунке отдельные точки соединены гладкой кривой для лучшей визуализации. Сплошной линией без дополнительных пометок на рис. 1 показана парабола, наилучшим образом приближающая экспериментально полученные значения. При этом три параметра параболы рассчитывались по представленным выше формулам.

Построенная таким образом непрерывная кривая с достаточной точностью приближает экспериментальные данные. На реальных данных по пациентам с СГМ величина $min\ J(a,b,c)$ практически никогда не превышала 0,02.

Представленный подход позволяет значительно сократить число параметров, описывающих восприятие пациентом различных временных отрезков, а получаемые при этом параметры a, b, c вместе с величиной квадратичного отклонения параболического приближения от экспериментальных данных в гораздо большей степени подходят для использования в информационной базе, отражая важный для агрегирующих параметров интегральный эффект. С другой стороны, непрерывная кривая значительно проще воспринимается исследователем, как характеристика состояния пациента, что найдет отражение в приводимых в дальнейшем примерах.

Другая возможность существенного снижения числа параметров, характеризующих исходное и текущее состояние пациента, может быть связана с автоматизацией расчета по известной методике [3] характерного для каждого пациента признака лево- или праворукости по физическим или функциональным особенностям.

В целом, уменьшение числа параметров, характеризующих состояние пациента, за счет агрегирования позволяет упростить анализ и использование такого рода данных.

Систематика психических расстройств острого периода СГМ

Собранные массивы данных о результатах обследований были предварительно систематизированы и затем сгруппированы по выделенным исследователями четырнадцати синдромам, что позволило сформировать таблицу характерных для каждого синдрома проявляющихся у всех пациентов группы симптомов, выделенных при первом обследовании, и, таким образом, получить набор значений отдельных параметров обследований, характерных для различного развития заболевания. При этом для дискретных параметров как совпадение рассматривалась полная идентичность, а для непрерывных – попадание в область с 2% относительным отклонением от среднего значения по группе.

Выделенные таким образом группы присутствующих симптомов являются несовпадающими и могут рассматриваться как базовые для перечисленных синдромов

Таблица количества симптомов, совпадающих у всех пациентов, отнесенных к одному синдрому

N N	Название синдрома	Количество совпадающих симптомов в векторе
1	гипоманиакальный продуктивный	7
2	гипоманиакальный непродуктивный	11
3	тревожной депрессии (на факт травмы)	10
4	тревожной депрессии (с НУВ)	3
5	тревожной депрессии с вновь возникающим поводом	4
6	истерической депрессии	14
7	апатической депрессии	14
8	тоскливой депрессии	20
9	астенической депрессии	7
10	анозогностический	6
11	дисмнестический	10
12	астенический (с явлениями дереализации и деперсонализации)	12
13	гиперастенический	18
14	гипоастенический	15

С другой стороны, выделение групп пациентов с совпадающими синдромами позволило характеризовать закономерности в векторах параметров $\vec{x}$ по каждой группе.

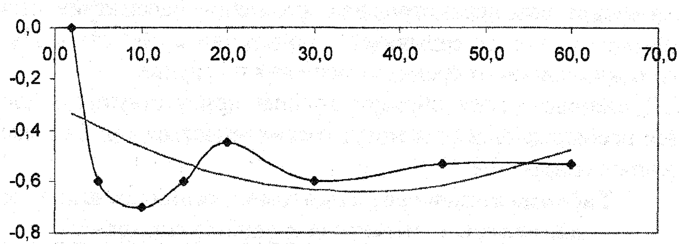


Рис. 2. Относительное отклонение оценок временных интервалов при тоскливой депрессии

Так, для группы пациентов с тоскливой депрессией можно заметить следующую закономерность в результатах первого обследования по оценке и воспроизведению временных интервалов. Оценка предлагаемых временных интервалов больными, отнесенными к данной группе систематически занижалась. Характерный пример представлен на рис. 2. с аналогичным рис. 1 обозначением кривых.

В то же время для эксперимента по воспроизведению пациентами данной группы временных интервалов по принятой методике было характерно завышение результатов (рис. 3).

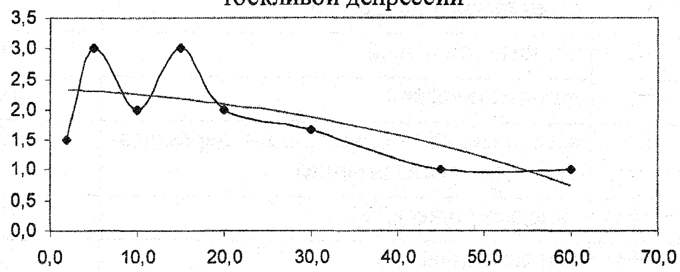


Рис. 3. Относительное отклонение воспроизведения временных интервалов при тоскливой депрессии

При указанном диагнозе подобная картина устойчиво сохранялась и в результатах последующих обследований на чувство времени.

Как пример с противоположным характером оценок по чувству времени можно выделить результаты в группе пациентов с гипоманиакальным продуктивным синдромом. Характерные зависимости для пациентов этой группы представлены на рис. 4 и рис. 5. При указанном синдроме данные наблюдений при последующих обследованиях плавно приходят в норму.

Рис. 4. Относительное отклонение оценок временных интервалов при продуктивном гипоманиакальном синдроме

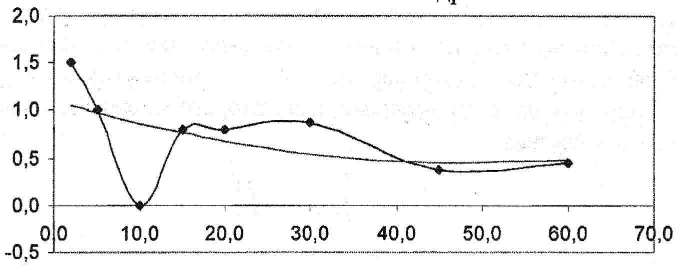
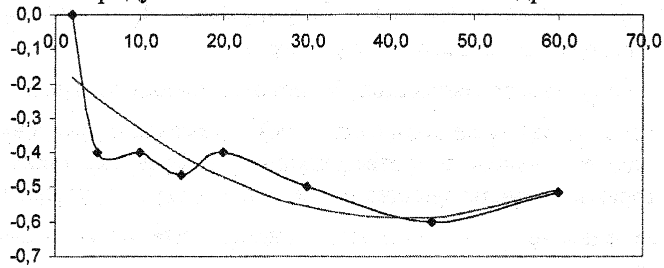


Рис. 5. Относительное отклонение воспроизведения временных интервалов при продуктивном гипоманиакальном синдроме



Отдельный интерес представляет исследование изменения данных от первого к последующим обследованиям, которое позволяет проследить развитие заболевания во времени. Для оценок подобной динамики аналогичным образом можно оценить изменение с течением времени суммарных (квадратичных) относительных отклонений индивидуальных параметров, полученных при обследовании пациентов, от нормативных значений. Однако, результаты в этой области заслуживают отдельного обсуждения.

Выделение аналогов по информационной базе

Наличие значительной информационной базы с результатами уже проведенных обследований пациентов с СГМ и уже определенным психиатрическим диагнозом может быть использовано при принятии более информативного решения по вновь поступившим пациентам за счет поиска наиболее близких (по значениям параметров) случаев среди представленных в базе.

Для этого может быть использована следующая процедура. Каждому поступающему пациенту ставится в соответствие получаемый при его первом обследовании вектор параметров $\bar{x}$, с подвергнутыми предварительной нормировке компонентами, для которого может быть рассчитана величина отклонения

$$h_j = \sum_{k=1}^{N} \left[x_k - x_k^j \right]^2 .$$

где $\bar{x}^j$ -- аналогичным образом нормированные векторы параметров первого обследования пациентов, имеющиеся в базе, $j=1,2,..., M$. Предварительная нормировка осуществлялась для того, чтобы в первом приближении уравнять вклад в вычисляемые нормы и невязки отдельных компонент вектора, и сводилась к нормировке максимальных возможных амплитуд A_k отдельных параметров. Как правило, $A_k = 2$, что характерно для большого числа задаваемых параметров.

Из получаемого множества M положительных величин h_j выделяется наперед заданное количество (m) наименьших значений. Параметры, представленные в соответствующих выбранным таким образом h_j векторах и соответствующие им пациенты могут рассматриваться как наиболее близкие по симптоматике аналоги для вновь обследуемого больного.

При этом для сокращения времени анализа всей информационной базы может быть сформирован стек из m элементов для хранения текущего набора пар $(j,\ h_j)$ с наименьшими текущими значениями величин h_j, упорядоченными по убыванию от корня стека. При вычислении значения h_j для очередного вектора параметров из информационной базы суммирование значения h_j можно проводить не для всего вектора, а только до тех пор, пока накапливаемая сумма не превысит значения h_m, находящегося в корне стека. Если же вновь полученная окончательная величина h_j окажется меньше хранящейся в корне, то она вытеснит из стека последнее значение и займет место в соответствии с расположением в стеке величин по убыванию от корня. При этом в предположении равномерного распределения величин h_j на некотором отрезке $[0,H]$, $H>0$, и равномерного распределения значений компонент по каждому отдельному вектору задача расстановки дополнительных текущих проверок с корневым значением при текущем суммировании эквивалентна задаче выделения m наименьших по норме элементов из массива, состоящего из M векторов $\vec{x}^j \in R^n$ с неотрицательными компонентами. Оптимальное расположение одной дополнительной проверки регламентируется следующим утверждением.

Пусть через $ET_0(n)$ обозначено математическое ожидание числа операций необходимых для суммирования компонент очередного n-го вектора и сравнения получаемой суммы с текущим корневым значением h_m в том случае, если дополнительное промежуточное сранение с величиной h_m не проводится. При этом для простоты будем считать, что операция одного суммирования двух чисел и одного сравнения двух чисел занимают равное время. Обозначим через $ET_{l(n)}(n)$ математическое ожидание числа операций такого же вида в случае, если проводится дополнительная проверка накапливаемой суммы компонент очередного n-го вектора после суммирования $l(n)$ элементов вектора. Пусть Z – множество целых чисел, $[z]$ -- целая часть числа z.

__Теорема.__ *Существует* $n \geq m\left(1+2\dfrac{1+\sqrt{N+1}}{N}\right)$ *такое, что*

$ET_{l(n)}(n) \leq ET_0(n)$ *при оптимально выбираемом месте дополнитеной проверки в виде*

$$\mathrm{l(n)}=\begin{cases}\sqrt{\dfrac{m(N+1)N}{n}}\in Z\\ \left[\sqrt{\dfrac{m(N+1)N}{n}}\right]\wedge\left[\sqrt{\dfrac{m(N+1)N}{n}}\right]+1\in Z\end{cases}$$

__Доказательство__. Для всех $n=1,2,\ldots,M$ величина $ET_0(n)=N$.

Для значения $s(l)$ частичной суммы $l(n)$ элементов текущего n-го вектора $\vec{x}^j \in R^n$ значения вероятностей имеют вид:

$$P(s(l) \le h_m) = \frac{mN}{nl}, \qquad P(s(l) > h_m) = \frac{nl - mN}{nl}.$$

Отсюда

$$ET_l(n) = (N+1)\ P(s(l) \le h_m) + l\,P(s(l) > h_m) = \frac{m(N+1)N + nl^2 - mNl}{nl} \Rightarrow \min_l$$

Исходя из необходимого условия экстремума в случае непрерывной зависимости от вещественного аргумента l можно получить значение оптимального $l(n)$, представленного в утверждении теоремы.

Подставив $l(n)$ в неравенство $ET_{l(n)}(n) \le ET_0(n)$, получим оценку снизу для текущего номера элемента, начиная с которого эффективно использовать дополнительную проверку с корневым значением h_m после суммирования ровно $l(n)$ компонент текущего n-го злемента массива.

Теорема доказана.

Заключение

Предлагаемая схема вычислительной обработки данных клинических обследований ориентирована на облегчение для пользователя восприятия значительных объемов имеющихся массивов информации и не должна восприниматься иначе, как вспомогательный инструмент для практикующего специалиста.

Авторы выражают искреннюю признательность сотрудникам НИИ нейрохирургии им. Н.Н.Бурденко к.м.н. Олегу Семеновичу Зайцеву и к.т.н. Евгению Леонидовичу Машерову за полезные консультации.

Литература

1. Непомнящий В.П., Лихтерман Л.Б., Ярцев В.В. Черепно-мозговая травма. Клиническое руководство. Том 1. М.: Антидор. 1998.
2. Васильев Ф.П. Методы решения экстремальных задач. М.: Наука, 1981.
3. Доброхотова Т.А., Брагина Н.Н. Левши. М.: Книга, 1994. 232 с.