РАСШИРЕНИЕ ВОЗМОЖНОСТЕЙ МЕТРИЧЕСКОГО ПОДХОДА НА ОСНОВЕ ТЕОРИИ СРЕДНИХ И ТЕОРИИ ОШИБОК

Сидельников Ю. В.

ИПУ РАН, МАИ

sidelnikovy@mail.ru

Аннотация. В материале предлагаются расширить возможности метрического подхода для решения специальных задач, например, в теории расписаний, ослабляя требования к аксиомам метрики или введение вероятностных мер близости. Рассматривается результаты автора на стыке теории средних и такой области исследования как показатели экспертных ошибок, которые аксиоматически заданы. Усилен результат, полученный акад. А.Н. Колмогоровым, при рассмотрении им системы аксиом для вывода аналитической формулы ассоциативной средней.

Ключевые слова: требования к аксиомам метрики; показатели экспертных оценок; теорема Колмогорова и ее усиление.

Введение

Метрики достаточно широко используются в фундаментальных и прикладных исследованиях. Например, в теории расписаний, которая является фундаментальной областью дискретной оптимизации [1].

Обзоры, в рамках которых рассматриваются всевозможные меры близости, публиковались достаточно давно. Например, обзор мер близости в пространстве множеств был опубликован Г. В. Раушенбахом в 1982 году [2**]**. В настоящее время появляются исследования, где рассматриваются принципиально новые меры близости (сходства), например, вероятностные меры близости, которые рассмотрены в монографии Д. З. Уздина [3**].** Полагаем, что метрический подход для решения разнообразных задач может получить новый импульс при использовании, не только мер близости, но и функций, которые играют роль метрики, однако метриками в классическом понимании не являются, так как их аксиомы «ослаблены», а также не дискретных, а вероятностных мер близости. И, такие работы есть, но их существенно меньше.

При этом, мы рассматриваем меру близости как однозначную, неотрицательную действительную функцию ρ(x,y), определенную для любых x и y на X и удовлетворяющую трем стандартным условиям (аксиомам) [стр. 48, 4].

Постановка задачи. Рассмотреть варианты перспективных направлений в фундаментальных и прикладных исследованиях, связанные с использованием вероятностного подхода, для мер близости и ослаблением требований к аксиоматике метрики, получив в их рамках новые результаты.

1 Возможные варианты перспективных направлений

Среди первых трех, по числу аксиом меры близости, вариантов перспективных направлений, мы рассмотрим лишь те, которые связаны с ослаблением требований к аксиоматике метрики.

Причем, если первое и второе лишь обозначим, то в третьем направлении попытаемся не только его предложить, но и получить в его рамках новые результаты.

В качестве первого направления предложим использовать для анализа информации не только метрики, но и псевдометрики[24]. В этом случае, ослабляется первая аксиома метрики (рефлексивность), таким образом, что не требуется выполнение следующего условия: чтобы из E (x,y) = 0 следовало бы x = y [Стр. 740-741, 7].

Мы исходим из следующего определения псевдометрического пространства как множества X, наделенного псевдометрикой E (x,y). При этом любым двум элементам x, y ϵ X ставиться в соответствие действительное число E такое, что:

первая аксиома: если x=y, то E (x,y) = 0, и аксиома: неравенство треугольника. Когда, для любой тройки x, y, z ϵ X , E (x,y) ≤ E (x,z) + E (z,y).

Легко показать, что псевдометрика симметрична и положительно полуопределена, т.е. E (x,y) > или = 0. (См., например, стр.31, [**5**]).

В качестве второго направления предложим рассмотреть вариант ослабления аксиомы неравенства треугольника. Как, например, в диссертации Г.В. Иофиной **[6].** Эта диссертация направлена на поиск способов измерения расстояний на объектах, задаваемых наборами порядковых признаков**.** Автор

[24] *В более ранних работах на русском языке данное понятие именовалось как: «квазирасстояние» [стр. 31, 5]. Встречаются работы, где данное понятие называется полуметрикой [6**.].***

диссертации полагает, что «иногда полезно рассматривать функции на значениях признаков, которые играют роль метрики, однако метриками в классическом понимании не являются». Цель работы этой диссертации — поиск и использование оптимальных функций расстояний, удовлетворяющих всем аксиомам метрик или полуметрик, кроме неравенства треугольника. Автор использует условие, замещающее неравенство треугольника, на более слабое — расстояние между дальними значениям не меньше, чем между ближними (условие порядка). Точнее, на каждом признаке задана своя функция расстояния, которая удовлетворяет всем аксиомам метрики кроме неравенства треугольника. То есть функция $p\ (x,\ y)$: $E^N\ x\ E^N \rightarrow E^M$ удовлетворяет следующим условиям:

$p\ (x, y) = 0 \Longleftrightarrow x = y$, $p\ (x, y) = p\ (y, x)$,

Если $x \geq y$, то $p\ (x, z) \geq p\ (y, z)$ и $p\ (x, z) \geq p\ (x, y)$, для любых z принадлежащих E^N, таких, что $z \leq y$.

При этом требуется, чтобы выбранная мера близости была оптимальна для различных задач, главным образом, задач распознавания. Кроме того, автор диссертации обосновано полагает, что результат работы алгоритмов классификации и кластеризации сильно зависит от метрик, заданных на объектах. И, значит, особо важно выбрать правильную функцию близости в задачах кластеризации [8, 9].

Третье перспективное направление. В данном случае, мы рассматриваем вариант, когда допускается асимметричность метрики и, при этом, ослабляются, тем или иным способом, первая и (или) третья аксиома метрики. В таком случае, можно использовать асимметричные меры, например, для числового случая показатель ошибки $E_1\ (x, y) = |x-y| \ / \ |y|$. Обычно все такие меры называют показателями ошибки. Использование таких асимметричных мер может быть, возможно, когда на множестве можно задать тернарное отношение «между». И полезно, когда нужно рассматривать качественно другие ситуации. Например, вы пришли за одну минуту до отхода поезда или опоздали на ту же минуту. Надо ли учитывать асимметричность меры в таких ситуациях. Несомненно, и, в дальнейшем, мы это покажем.

Четвертое перспективное направление — введение и использование вероятностных и вероятностно-метрических мер близости и их эффективный подбор.

2 Обзор существующих современных публикаций по данной тематике

Рассмотрим, в первую очередь, обзор публикаций, в рамках которых используются асимметричные меры близости. Таких работ, по сравнению с исследованиями с использованием обычных метрик, гораздо меньше.

Приведем ряд примеров с использованием асимметричных мер. Необходимость учета асимметричности меры возникает, например, при формализации понятия степени приближения к истине отдельных фрагментов научного знания — теорий, гипотез, научных утверждений и т.п. в рамках решения проблемы правдоподобности научного знания (логической теории правдоподобности). Пример, иллюстрирующий сложности этой задачи рассмотрен в монографии основателя этой теории Карла Поппера, которая базируются на основе аппарата формальной логики. В качестве пояснения он рассмотрел ситуацию с опозданием на поезд, а для учета степени приближения (меру правдоподобности) к истине использует показатель $E\ (x,\ y) = |x-y|$, задавая затем риторический вопрос: «Отклонения от истинного времени в обоих случаях равны, но равная ли мера правдоподобности должна быть приписана этим утверждениям? На основе только одной интуиции решить этот вопрос нельзя, и, во всяком случае, во многих практических ситуациях (например, отправление поезда) ошибка одного существенно менее значима, чем ошибка другого» [10].

Следующий пример употребления асимметричных мер можно обнаружить в исследованиях по кластерному анализу, который используется в различных научных и прикладных областях и также является актуальной темой исследований. Так, в работе А. Р. Айдиняна и О. Л. Цветковой, в отличие от большинства существующих работ предложены алгоритмы, предназначенные для кластеризации объектов, которые описываются векторами признаков в пространстве с несоблюдением аксиомы симметрии [11]. «Суть первого из предложенных алгоритмов кластеризации заключается в последовательном формировании кластеров с одновременным перенесением кластеризованных объектов из ранее созданных кластеров в текущий кластер в случае, если это уменьшит критерий качества. По сравнению с существующими алгоритмами неиерархической кластеризации такой подход к формированию кластеров позволяет уменьшить вычислительные затраты» [11].

По мнению исследователей из ИПУ РАН: «недостатком большинства мер, основанных на онтологических структурах, является симметричность (экспертные оценки показывают, что мера близости не всегда симметрична). Кроме того, эти меры независимы от контекста и чувствительны к структуре иерархии» [12].

Среди англоязычных работ можно отметить следующие:

- В материале [13]: «предлагается асимметричная мера семантической близости. В зависимости от направления прохождения ребрам придается разный вес, так как потомок более подобен родителю, чем родитель потомку».
- В статье [14] представлена концепция асимметричной модели сходства и взвешенного вектора частот слов или термов (HFV) и сконструированы две новые асимметричные меры: HFM и HIPM. Авторы показали, что симметричные меры сходства не всегда могут эффективно работать в рамках общей проблемы измерения сходства текстов. Будь то, информационной поиск, интеллектуальный анализ текстов, извлечение веб-контента, классификация и кластеризация текстов в коллекции и обнаружении копии документов в смысле плагиата и т. д.. Такого рода направления исследования требуют методов способных обрабатывать неструктурированную информацию, поэтому самым популярным подходом является схема, основанная на частоте повторений того или иного элемента, в случае текстов, слов или сочетаний символов. Для документов эта схема использует вектор частот слов из словаря, построенного на всех изучаемых документах, тем самым обеспечивается единая размерность. А стандартными векторными мерами сходства являются функция косинуса, скалярное произведение и функция пропорций. Но все они симметричны.
- Представляют интерес и другие англоязычные исследования, см. например: [15, 16, 17].

Обзор исследований с использованием ассиметричной меры близости показал, что хотя такие работы и существуют, но и в них не учитывается ряд трудностей их использования. Рассмотрим их.

3 Трудности использования асимметричных мер близости

Первая трудность. Мы не можем использовать вышеуказанное определение псевдометрического пространства, так как из него вытекает именно симметричность меры. Таким образом, необходимо либо:

1. Ослаблять первую аксиому и (или) аксиому: «неравенство треугольника»;
2. Использовать нижеследующие аксиоматики для показателей ошибки, либо:

«Определение 3.1. Показателем ошибки E называется действительная функция E = E(x,y), определенная на $X^N \times X^N$ и удовлетворяющая следующим условиям:

1. $E(x,y) = 0 \Longleftrightarrow x = y$;
2. $E(x,y) \geq 0$;
3. Для каждого фиксированного $x \in X^N$ функция $E_x(y) = E(x,y)$ строго монотонно убывает на луче $\{y \in X^N: y<x\}$ и монотонно возрастает на луче $\{y \in X^N: y>x\}$;
4. Для каждого фиксированного $y \in X^N$ функция $E_y(x) = E(x,y)$ строго монотонно убывает на луче $\{x \in X^N: x<y\}$ и строго монотонно возрастает на луче $\{x \in X^N: x>y\}$». (См. стр. 170, [18]), либо:

«Определение 3.4. Показателем ошибки E называется любая действительная функция E(x,y), определенная на X×У и удовлетворяющая следующим условиям:

1. $E(x, y)=0 \Longleftrightarrow x = y$;
2. $E(x, y) \geq 0$;
3. Для любого фиксированного $x \in X$, функция $E_x(y)$ строго монотонно убывает на луче $\{y \in Y: y<x\}$ и монотонно возрастает на луче $\{y \in Y: y>x\}$;
4. Для каждого фиксированного $y \in Y$, функция $E_y(x)$ строго монотонно убывает на луче $\{x \in X: x<y\}$ и строго монотонно возрастает на луче $\{x \in X: x>y\}$;
5. $E(x, y)$ непрерывно по y».

При этом Y — линейно упорядоченное отношением ($\leq$) множество, наделенное порядковой топологией, а X — связное подмножество Y. (См. стр. 173 [18]).

Заметим, что, второе определение показателя ошибки отличается от первого не только областью определения, но и наличием пятого условия, которое позволяет рассмотреть новый необходимый и достаточный критерий.

Вторая трудность. Необходимо учитывать, что вывод при сопоставительном анализе, полученный на основе используемой ассиметричной меры, (показатель ошибки) может зависеть от вида, точнее от класса эквивалентности, использованной меры. Возможно, первый, кто обратил на это внимание, был Г. Галилей. В его письме к Ноццолино, содержащей в монографии К. Джини, рассмотрен следующий условный пример [19]: «Пусть перед покупкой лошака один знаток оценил скотину в 10 скудо (денежная единица того времени в Италии), второй - в 1 тыс. скудо, а хозяин продал лошака за 100 скудо. Спрашивается, какая из оценок точнее? Ноццолино утверждал, что второй знаток допустил большую ошибку, потому что его оценка отличается от фактической на 900 скудо, тогда как оценка первого - лишь на 90. На что Галилей, возразил, что превышение в оценке второго эксперта равняется преуменьшению в оценке первым, и, значит, ошибки одинаковы».

Таким образом, перед отбором меры необходимо обосновать, какую меру выбрать, а точнее, какой класс эквивалентности таких мер нам подходит. Приведем поясняющий пример. Пусть необходимо выяснить, подойдет ли для верификации прогноза прироста ВВП РФ на 2021 год следующая мера (показатель ошибки) E6(x, y) = $|\ln x/y|$, где x, y $\in$ R, x — экспертная оценка, y — истинное значение. Первый эксперт дал прогнозную оценку X1 = 1% ожидаемого прироста, второй эксперт дал оценку X2= 25% ожидаемого прироста ВВП по сравнению с 2020 годом. Подводя итог 2021 года, экономисты получили оценку прироста ВВП по сравнению с 2020 годом пять процентов (Y= 5 %). Если мы воспользуемся логарифмическим показателем ошибки, то получим, что ошибки двух прогнозов одинаковы. ($|\ln 1/5| = |\ln 25/5|$). На наш взгляд, такой показатель для верификации прогноза прироста ВВП РФ на 2021 год не подходит. И в следующий раз нам не нужно приглашать для оценивания второго эксперта. Значит, для каждой ситуации надо находить подходящий показатель ошибки. Такой пример является подсказкой для нахождения процедуры подбора вида показателя ошибки, точнее, класса эквивалентности показателей ошибки.

Именно с этим и связана третья трудность: построение тех или иных вариантов процедуры подбора вида показателя ошибки. В монографии автора предложен один из вариантов процедуры подбора класса эквивалентности видов показателей ошибки на основе экспертных заключений [стр. 189, 18]. Поясним последнее утверждение. Классы эквивалентности показателей ошибки вполне естественно вводятся в рамках:

- любых двух различных аксиоматических заданий показателей ошибки: определение 3.1. [стр. 170, 18] и определение 3.4. [стр. 173, 18];
- отношения знаковой эквивалентности (~). См. определения 3.3 [стр. 171, 18];
- Утверждения 3.1. «Отношение знаковой эквивалентности (~) есть отношение эквивалентности» [стр. 171, 18].

В рамках рассмотренной и вновь создаваемых аналогичных экспертных процедур возникает необходимость усреднения экспертных оценок. Возникают следующие задачи:

1. Вид используемой средней величины необходимо обосновывать.
2. Кроме того, необходимо уметь рассчитывать найденный вид средних. Но, для этого, необходимо чтобы они были заданы не аксиоматически, а в аналитическом виде.

Поясним последнюю задачу. В случае, если основой для отбора вида средней является следующая аксиома: «можно заменить некоторую группу значений их собственным средним, не меняя общего среднего», то нужно обратить внимание на результат ак. А.Н. Колмогорова [20, 21]. В этой работе А.Н. Колмогорова было показано, что каждый тип среднего M*, если он удовлетворяет четырем условиям:

1. M*(x1,x2,…….,xn) – непрерывна и монотонна по каждому переменному (для определенности будем считать, что M*– возрастает по каждому переменному);
2. M*(x1,x2,…….,xn) – симметрическая функция;
3. Среднее от одинаковых чисел равно их общему значению: M*(x,x,…….,x) = x;

Можно заменить некоторую группу значений их собственным средним, не меняя общего среднего: $M^*(x_1,x_2,\ldots\ldots,x_m,\ y_1,y_2,\ldots,y_n) = M^*_{n+m}(x,\ x,\ldots,\ x,\ y_1,y_2,\ldots,y_n)$, где $x=M^*(x_1,\ x_2\ldots,\ x_m)$ необходимо имеет вид: $M^*(x_1,x_2,\ldots\ldots,x_n) = F^{-1}(F(x_1) + F(x_2)+\ldots+ F(x_n)/n)$, где F – непрерывная строго монотонная функция, ($x_i \in [a,b] \subset R$).

Таким образом, мы имеем возможность рассчитывать значения такой средней, на основе аналитической формулы. Но, в нашем случае, а мы включаем и асимметричные меры, а вторая аксиома это запрещает. Таким образом, необходимо получить более общую формулу для любой ассоциативной средней, но без второй аксиомы.

В исследовании де Финетти показано, что любая ассоциативная средняя является монотонной [22], таким образом, ему удалось усилить результат ак. А.Н. Колмогорова.

Нашей задача состояла в ослаблении требования симметричности.

В том случае, если мы находим среднюю величину для множества, без учета его формирования, условие симметричности является вполне естественным, но уже для множества, которое формируется из последовательности, по-видимому, нет. Для подтверждения последнего рассмотрим такой широко распространенный способ формирования экспертных групп как метод снежного кома [стр. 159, 18]. В этом случае, вполне уместно полагать, что эксперт, называя другого эксперта, вряд ли назовет того, у которого уровень знаний по задаваемому вопросу будет ниже, чем у него самого. Таким образом, вполне уместно назначать большую весомость числовой оценки последующего эксперта.

Общая постановка задачи: предложить аналитическое выражение для любой ассоциативной средней, но без условия симметричности, частным случаем которой была бы формула Колмогорова.

Теорема[25]. Для любой строго монотонной и непрерывной действительнозначной функции F: $R^0 \to R$ следующее выражение является формулой для непрерывной ассоциативной средней:

M*(x1,x2,…….,xn) = F-1 {Σ1n qi F(xi)},

где $x_i \in R^0$– произвольному ограниченному интервалу действительных чисел, $\Sigma_1^n q_i = 1$, $q_i \geq 0$, $q_i \in R$.

2 Направления будущих исследований в сфере расширения возможностей метрического подхода

Предлагается рассмотреть следующий вариант расширения метрического подхода. Необходимо обратить внимание, что во всех вышеуказанных случаев, областью значения мер близости, или функций, которые играют роль метрики, является положительно определенное подмножество действительных чисел.

На наш взгляд, желательно расширять трактовку и этого элемента в аксиоматике меры близости. Для того рассмотрим следующие объекты, на множестве которых полезно ввести аналоги мер близости и соответствующее понятие «близость».

Среди них, например, такие объекты как: меры близости.

Поясним полезность введения мер близости на мерах близости.

Именно для них могут быть полезны различные варианты изменения области значения мер близости, или функций, которые играют роль метрики.

Предварительно необходимо отметить, что в рамках различных исследований, иногда их авторы ставят вопрос относительно наиболее подходящих мерах близости на множествах различной природы. Например, в диссертации Г.В. Иофиной, автор требует, чтобы выбранная мера близости была оптимальна для различных задач, главным образом, задач распознавания [6**].**

В монографиях Н. Г. Загоруйко и В. Д. Мазурова говорится о выборе правильной функции близости в задачах кластеризации [8, 9].

Поиск наиболее подходящих мер близости обусловлен различиями реальных и виртуальных объектов, в математических моделях которых используются эти меры близости. Но, можно и нужно пойти дальше.

Итак, в дальнейшем исследовании, мы исходим из следующего.

Постулата №1

В рамках одного и того же реального или виртуального объекта, можно, а в ряде случаев необходимо, использовать различные меры близости или их расширения как инструмент их анализа.

При этом мы исходим из того, что эти объекты могут иметь неоднородную структуру, требующие для своего изучения различные меры в различных частях этих же объектов.

Приведем поясняющий пример. Для этого рассмотрим как множества, которые обладают как одинаковыми свойствами во всех направлениях — изотропные, так и анизотропные, т. е. обладающими различными свойствами в разных направлениях.

Рассмотрим анизотропные множества, которые содержат *N* не пустых изотропных подмножеств.

При этом, эти множества могут быть отображением как реальных, так и виртуальных объектов.

Постулат № 2.

Полагаем, что для исследования не только изотропных, но и анизотропных непустых множеств необходимо, на этих множествах, ввести понятия направления.

---

[25] *Данная теорема была впервые доказана в диссертации автора и анонсирована в его автореферате [23].*

Это возможно, например, введя линейный или частичный порядок на множестве элементов этих множеств, а для дальнейших исследований, ввести и меру близости или ее различные расширения на элементах этих множеств.

Но, рассматривая различные меры близости на одном и том же объекте, необходимо исследовать как преобразуются эти меры при переходе из одного изотропного множества в другое, в рамках одного и того же объекта. Таким образом, желательно ввести аналог меры близости на мерах близости — «супермеру близости».

Введем понятие супермеры близости *(P)* на множестве мер близости *E(x,y),* и исследуем возможности такой метрической конструкции — $P\ \{E1(x,y),\ E2(x,y)\},\ x,y \in X,\ i= 1,2,,,,n.$

Полагаем, что на множестве мер близости *E(x,y),* можно ввести как линейный, так и частичный порядок.

Кроме того, для этого отображения необходимо рассмотреть области определения и значения однозначного отображения *P* (*E(x,y)*), а также ее закон соответствия. Кроме того, ввести понятия «различия» и «совпадения» значения супермеры на множестве мер близости.

Полагаем, что область определения отображения *P* (*E(x,y)*), есть все возможные меры близости или их расширения, например, псевдометрика.

На область определения (множество мер близости) можно ввести отношение линейного или частичного порядка, а также ввести тернарное отношение «между».

Это возможно, например, на основе следующего Утверждения 3.16. автора: «Пусть У — линейно упорядоченное отношением (≤) множество, наделенное порядковой топологией[1], а *X* - связное подмножество *У*, *Ф* - непрерывная строго монотонно возрастающая действительнозначная функция. Тогда функция

$$E(x,y) = \begin{cases} K1\ [\Phi(x)-\Phi(y)], & \text{если } x>y, \\ K2\ [\Phi(y)-\Phi(x)], & \text{если } y \geq x, \end{cases}$$

является показателем ошибки, где $x \in X;\ y \in Y;\ K1, K2 \in R+;\ K1 \neq K2$ » [См. стр. 188, 18].

Так как *K1* и *K2* являются различными действительными числами, то при фиксированной функции Ф, они определяют различные меры близости. А отношение линейного порядка на действительной прямой индуцирует линейный порядок на множестве мер близости заданных в Утверждении 3.16.

Полагаем, что понятие «совпадения» между мерами близости на множестве мер близости можно ввести, например, через отношение «тождества», между функциями $E^i(x,y)$.

Необходимо обратить внимание, что область значения супермеры является множеством преобразований одной из мер в другую. В случае дискретного множества преобразований, её можно задать как минимальное количество преобразований необходимых для перехода от одной из мер в другую и таким образом получить положительное число.

Пояснить понятие «преобразование» одной из мер в другую можно, следующим образом. Пусть, например, имеем два показателя ошибок $E(x.y)= |x - y|$, и $E\ (x.y)= |ln_n\ x - ln_n\ y|$. Рассмотрим, как преобразуется одна из мер близости, выраженной функцией E(x.y), в другую функцию $E`(x.y)$.

Полагаем, что «переход» от одной из функции в другую, в этом конкретном случае, возможен можно через преобразование $ln_n x$ в x. График функции $z=ln_n x$, при $x \geq 1$ приближается к графику функции $z=x$ при стремлении основания логарифма к бесконечности.

Мера различия между мерами близости на множестве мер близости может выражаться в порядковой шкале, например, в баллах. Приведем поясняющий пример, такого различия. Рассмотрим два различных токаря, но их различие будет только в их разрядах. Один из них, будет токарем третьего разряда, а второй токарь — второго разряда.

Покажем полезность и необходимость такой конструкции. При этом будем исходить из Постулат ***№ 3.***

Само множество, где определены меры близости *E(x,y)* может быть не только изотропным, т. е. обладает одинаковыми свойствами во всех направлениях, но и анизотропным, т. е. обладает различными свойствами в разных направлениях.

Так, например, в реальном космическом пространстве вблизи массивных тел для расчетов используется метрика Минковского, а вдали от массивных тел можно использовать метрику Эвклида.

Порядковой топологией называется наименьшая система подмножеств множества ***Y,*** *замкнутая относительно объединения и конечного пересечения, и содержащая все открытые лучи [24]****.***

На наш взгляд, необходимо исследовать свойства такой метрической конструкции, где введена супермера близости *(P)* на множестве мер близости Ei.(x,y).

Заключение

В материале сначала рассмотрены три варианта расширения метрического подхода для решения специальных задач, например, в теории расписаний, связанные с ослаблением требований к аксиоматике метрики. Кроме того, предложено поддержать вариант расширения метрического подхода, на основе вероятностных мер близости и вероятностно-метрических мер близости рассмотренных в монографии Д. З. Уздина [3].

Указано на перспективность выбора функций близости для разнообразных типов задачах. Так, в исследовании Г.В. Иофиной, автор требует, чтобы выбранная мера близости была оптимальна для различных задач, главным образом, задач распознавания. Н. Г. Загоруйко и В. Д. Мазуров говорят о выборе правильной функции близости в задачах кластеризации [8, 9]. Но, как выбирать оптимальную или правильную меру близости? Этот вопрос пока не решен в полной мере. Хотя автору, для асимметричных числовых мер близости удалось построить такую процедуру подбора вида показателя ошибки. (См., например, стр. 189 [18]).

Кратко описаны результаты автора, связывающие наиболее известные аналитические средние с характеристиками классов эквивалентности числовых показателей ошибки, в рамках двух различных аксиоматических заданий показателей ошибки.

Усилен результат, полученный А. Н. Колмогоровым, при рассмотрении им системы аксиом для вывода аналитической формулы ассоциативной средней. В работе была получена формула для ассоциативной средней, но уже без условия ее симметричности.

Предложено новое направление будущих исследований в сфере расширения возможностей метрического подхода, связанное с расширением метрического подхода. Введено понятие супермеры близости и рассмотрен случай, когда областью определения супермеры близости являются меры близости, или функции, которые играют роль метрики. Показано полезность данной конструкции.

Приложение

Доказательство теоремы.

Во-первых, докажем, что $M^*(x_1,x_{2,\ldots\ldots,}x_n)$ является – средним по Коши, т.е.

min xi ≤ M*(x1,x2,…….,xn) ≤ max xi

Действительно пусть F - строго монотонно возрастающая функция, тогда

Σ1n qi F(min xi) ≤ Σ1n qi F(xi) ≤ Σ1n qi F(max xi),

F(min xi) ≤ Σ1n qi F(xi) ≤ F(max xi),

Аналогично можно показать и для F – строго монотонно убывающей функции.

Теперь докажем, что средняя, выраженная в виде $M^*(x_1,x_2,\ldots\ldots,x_n) = F^{-1}\{\Sigma_1^n q_i F(x_i)\}$, является ассоциативной.

То есть $M^*(x_1,x_2,\ldots\ldots,x_m, x_{m+1},\ldots,x_n) = = M^*_n (x,x,\ldots\ldots,x, x_{m+1},\ldots,x_n)$,

где $x=M^*_m(x_1,x_{2,\ldots\ldots,}x_m)$ для любого m< n.

Формула для среднего при m<n, образуется следующим образом $x=M^*_m(x_1,x_{2,\ldots\ldots,}x_m) = F^{-1}\{ \Sigma_1^m (q_i / \Sigma_1^m q_i) F(x_i)\}$, то есть, сохраняя прежние коэффициенты, мы заново нормируем их, с тем, чтобы сохранить условие

Σ1m qi *= Σ1m (qi / Σ1m qj) = 1.

M*n (x,x, xm+1,…….,xn) = F-1 { Σ1m qi F(x) + Σm+1nqi F(xi)}=

=F-1{Σ1m q i F [F-1 (Σ1m (qi / Σ1m qj) F(xi))] + Σm+1nqi F(xi)} =

= F-1{Σ1m q i [Σ1m (qi / Σ1m qj) F(xi)] + Σm+1nq i F(xi)} =

= F-1{Σ1mqi F(xi) + Σm+1nqi F(xi)} = M* (x1,x2, ….,xn).

Теорема доказана.

Литература

1. *Alexander A. Lazarev, Darya V. Lemtyuzhnikova, Frank Werner.* A metric approach for scheduling problems with minimizing the maximum penalty// Applied Mathematical Modelling. Oxford, Great Britain, Elsevier. 2021, № 89. P. 1163-1176.
2. *Раушенбах Г.В.* Меры близости в пространстве множеств// Алгоритмы анализа данных социально-экономических исследований/ ЭИ и ОПП СО АН СССР. - Новосибирск, 1982. С. 29-44.
3. *Уздин Д. З.* Новые меры близости, функции состояний и решающие правила в теории распознавания состояний (статистической классификации). — 2-е изд., доп. и испр. М.: МАКС Пресс, 2016. — 128 с.
4. *Колмогоров А. Н., Фомин С.В.* Элементы теории функций и функционального анализа. – М.: Наука, 1968. – 496 с.
5. Коллатц Л. Функциональный анализ и вычислительная математика. Перевод с немецкого И.Г. Нидеккера, под редакцией А.Д. Горбунова. М.: Изд-во «Мир», 1964. — 448 с.
6. Иофина Г. В. Выбор оптимальных метрик в задачах распознавания с порядковыми признаками». Диссертации на соискание кандидата физико-математических наук: 05.13.17 – Теоретические основы информатики. Москва. 2010. – 105 с.
7. Математическая энциклопедия. — М.: Советская энциклопедия. Т. 4. Ок — Сло. 1984. С. 1216.
8. *Загоруйко Н. Г.* Прикладные методы анализа данных и знаний. Новосибирск: Изд. Института математики СО РАН, 1999. — 270 с.
9. *Мазуров В. Д.* Метод комитетов в задачах оптимизации и классификации. М.: Наука, 1990. — 248 с.
10. *Popper K. R.* Conjectures and Refutations. The Growth of Scientific Knowledge. Ch. 10, section 3. London: Routledge and Kegan Paul; N. Y. Basic Books Inc., 1963.
11. *Айдинян А. Р., Цветкова О. Л.* Алгоритмы кластерного анализа задач с асимметричной мерой близости // Сибирский журнал вычислительной математики. 2018. № 2. С. 127 – 138.
12. *Крюков К.В., Панкова Л.А., Пронина В.А., Шипилина Л.Б.* Меры семантической близости в онтологиях // Проблемы управления. 2010. № 5. С. 2–14.
13. *Henrik Bulskov, Rasmus Knappe, Troels Andreasen*. On Measuring Similarity for Conceptual Querying. In T. Andreasen, A. Motro, H. Christiansen, H.L. Larsen (Eds.): Flexible Query Answering Systems, Lecture Notes in Artificial Intelligence Berlin: Springer, 2002. V. 2522. P. 100–111.
14. *Jun-Peng Bao, Jun-Yi Shen, Xiao-Dong Liu, Hai-Yan Liu*. Quick asymmetric text similarity measures // Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2-5 November 2003. Т. 1. Р. 374-379.
15. *Hubert L. J.* Min and max hierarchical clustering using asymmetric similarity measures // Psychometrika. 1973. 38. P. 63-72.
16. *Faith D. P.* Asymmetric binary similarity measures // Oecologia. 1983. Т. 57. №. 3. Р. 287-290.
17. *Chen H. H., Giles C. L.* ASCOS++ An Asymmetric Similarity Measure for Weighted Networks to Address the Problem of SimRank //ACM Transactions on Knowledge Discovery from Data (TKDD). 2015. Т. 10. №. 2. Р 1-26.
18. *Сидельников Ю. В.* Системный анализ технологии экспертного прогнозирования. М.: Изд-во МАИ-ПРИНТ, 2007. — 348 с.
19. *Джини К.* Средние величины. М.: Статистика, 1970.
20. *Kolmogorov, A. N.,* Sur la notion de moyenne. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur., (6) 12 (1930), 388–391.
21. *Колмогоров А. Н.* Об определении среднего / Математика и механика // Избранные труды / отв. ред. С. М. Никольский, сост. В. М. Тихомиров. — М.: Наука, 1985. — Т. 1. — С. 136-138.
22. *Finetti B. de.* Sul concetto di media // Giornale dell Instituto Ital. degli Attuari, 1931, Anno 11, No. 3, P. 365 - 396.
23. *Сидельников Ю. В.* Автореферат на соискание ученой степени доктора технических наук «Технология экспертного прогнозирования» Москва, 25 апреля, 2002, ОКОРОИ ИМЭМО РАН.
24. *Пфанцагель И.* Теория измерений. - М.: Мир, 1976.