



Math-Net.Ru

Общероссийский математический портал

В. Г. Горский, А. А. Гриценко, А. И. Орлов, Метод согласования кластеризованных ранжировок,
Автомат. и телемех., 2000, выпуск 3, 159–167

<https://www.mathnet.ru/at258>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<https://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 109.252.119.48

27 июня 2025 г., 10:55:42



Моделирование поведения и интеллекта

УДК 330.115

© 2000 г. В. Г. ГОРСКИЙ, д-р техн. наук
(Государственный научный центр РФ "Государственный научно-исследовательский институт органической химии и технологии Москва),
А. А. ГРИЦЕНКО
(ЗАО "L-Card Москва),
А. И. ОРЛОВ, д-р техн. наук
(Московский государственный технический университет им. Н. Э. Баумана, Москва)

МЕТОД СОГЛАСОВАНИЯ КЛАСТЕРИЗОВАННЫХ РАНЖИРОВОК ¹

Рассматриваемая проблема состоит в выделении наиболее подробного частичного порядка, общего для набора кластеризованных ранжировок (на статистическом языке – ранжировок со связями). Этот набор может отражать мнения нескольких экспертов или быть получен при обработке мнений экспертов различными методами. Предлагается метод согласования кластеризованных ранжировок, позволяющий "загнать" противоречия внутрь специальным образом построенных кластеров (групп), в то время как упорядочение кластеров соответствует упорядочениям всех исходных кластеризованных ранжировок.

1. Введение

В различных прикладных областях возникает необходимость анализа нескольких кластеризованных ранжировок объектов.

К таким областям относятся технические исследования, экология, менеджмент, экономика, социология, прогнозирование и т.д., особенно те их разделы, что связаны с экспертными оценками (см., например, [1, 2]). В качестве объектов могут выступать образцы продукции, технологии, математические модели, проекты, кандидаты на должность и др. Кластеризованные ранжировки могут быть получены как с помощью экспертов, так и объективным путем, например, при сопоставлении математических моделей с экспериментальными данными с помощью того или иного критерия качества. Описанный в настоящей статье метод разработан в связи с проблемами химической безопасности биосферы [3, 4] и экологического страхования [5, 6].

Предлагается новый метод построения кластеризованной ранжировки, согласованной (в раскрытом ниже смысле) со всеми рассматриваемыми кластеризованными ранжировками. При этом противоречия между отдельными исходными ранжировками оказываются заключенными внутри кластеров согласованной ранжировки. В результате упорядоченность кластеров отражает общее мнение экспертов, точнее, то общее, что содержится в исходных ранжировках.

В кластеры заключены объекты, по поводу которых некоторые из исходных ранжировок противоречат друг другу. Для их упорядочения необходимо провести новые исследования. Эти исследования могут быть двух типов. Во-первых, осуществление дополнительной математической обработки экспертных данных (вычисление

¹ Работа выполнена при поддержке Международного научно-технического центра (проект №317) и Российского фонда фундаментальных исследований (проект №97-06-80033).

медианы Кемени, нахождение группового выбора с помощью упорядочения по средним рангам или по медианам и т.п. [1, 2]). Во-вторых, привлечение дополнительной информации из соответствующей прикладной области, возможно, проведение прикладных научных работ с целью сравнения объектов из одного кластера.

Введем необходимые понятия, затем впервые сформулируем алгоритм согласования в общем виде и рассмотрим его свойства.

2. Определение кластеризованной ранжировки

Пусть имеется конечное число объектов, которые мы для простоты изложения будем изображать натуральными числами $1, 2, 3, \dots, k$ и называть *носителем*. Под кластеризованной ранжировкой, определенной на заданном носителе, понимаем следующую математическую конструкцию. Пусть объекты разбиты на группы, которые будем называть кластерами. В кластере может быть и один элемент. Входящие в один кластер объекты будем заключать в фигурные скобки. Например, объекты $1, 2, 3, \dots, 10$ могут быть разбиты на 7 кластеров: $\{1\}$, $\{2, 3\}$, $\{4\}$, $\{5, 6, 7\}$, $\{8\}$, $\{9\}$, $\{10\}$. Кластеры не имеют общих элементов, а объединение их (как множеств) есть все рассматриваемое множество объектов.

Вторая составляющая кластеризованной ранжировки – это строгий линейный порядок между кластерами. Задано, какой из них первый, какой второй, и т.д. Будем изображать упорядоченность с помощью знака $<$. При этом кластеры, состоящие из одного элемента, будем для простоты изображать без фигурных скобок. Тогда кластеризованную ранжировку на основе введенных выше кластеров можно изобразить так:

$$A = [1 < \{2, 3\} < 4 < \{5, 6, 7\} < 8 < 9 < 10].$$

Конкретные кластеризованные ранжировки будем заключать в квадратные скобки. Если для простоты речи термин “кластер” применять только к кластеру не менее чем из 2-х элементов, то можно сказать, что в кластеризованную ранжировку A входят два кластера $\{2, 3\}$ и $\{5, 6, 7\}$ и 5 отдельных элементов.

Введенная описанным образом кластеризованная ранжировка является бинарным отношением на множестве $\{1, 2, 3, \dots, 10\}$. Его структура такова. Задано отношение эквивалентности с 7-ю классами эквивалентности, а именно, $\{2, 3\}$, $\{5, 6, 7\}$, а остальные состоят из оставшихся 5 отдельных элементов. Затем введен строгий линейный порядок между классами эквивалентности.

Введенный математический объект известен в математической статистике как “ранжировка со связями” (см., например, [7, с. 48 и др.]). Дж. Кемени и Дж. Снелл используют термин “упорядочение” [8, с. 20], Б. Г. Миркин – “квазисерия” [9, с. 37], Ю. А. Шрейдер – “совершенный квазипорядок” [10, с. 127, 130]. Учитывая разноречивость в терминологии, мы сочли полезным ввести собственный термин “кластеризованная ранжировка” поскольку в нем явным образом названы основные элементы изучаемого математического объекта – кластеры, рассматриваемые на этапе согласования ранжировок как классы эквивалентности, и ранжировка – строгий совершенный порядок между ними (в терминологии [10, гл. IV]).

3. Определение сильной противоречивости

Следующее важное понятие – противоречивость. Оно определяется для четверки – двух кластеризованных ранжировок на одном и том же носителе и двух различных объектов – элементов того же носителя. Мы будем рассматривать два вида противоречивости – сильную и слабую. Два элемента из одного кластера будем связывать символом равенства $=$ как эквивалентные.

Пусть A и B – две кластеризованные ранжировки. Пару объектов (a, b) назовем *сильно противоречивой* относительно A и B , если эти два элемента по-разному упорядочены в A и B , т.е. $a < b$ в A и $a > b$ в B либо $a > b$ в A и $a < b$ в B . Отметим, что в соответствии с этим определением пара объектов (a, b) , эквивалентная хотя бы в одной кластеризованной ранжировке, не может быть сильно противоречивой: эквивалентность $a = b$ не образует “сильного противоречия” ни с $a < b$, ни с $a > b$. Совокупность строго противоречивых пар объектов для двух кластеризованных ранжировок A и B назовем *ядром сильных противоречий* и обозначим $S(A, B)$.

В качестве примера рассмотрим три кластеризованные ранжировки: A (см. выше), B и C , где

$$B = \{[1, 2] < [3, 4, 5] < 6 < 7 < 9 < [8, 10]\},$$

$$C = [3 < [1, 4] < 2 < 6 < [5, 7, 8] < [9, 10]].$$

Они определены на одном и том же носителе $\{1, 2, 3, \dots, 10\}$. Как нетрудно проверить, ядра сильных противоречий таковы:

$$S(A, B) = [(8, 9)], \quad S(A, C) = [(1, 3), (2, 4)],$$

$$S(B, C) = [(1, 3), (2, 3), (2, 4), (5, 6), (8, 9)].$$

Как при ручном, так и при программном нахождении ядра можно в поисках сильно и слабо (см. ниже) противоречивых пар просматривать пары $(1, 2)$, $(1, 3)$, $(1, 4), \dots$, $(1, k)$, затем $(2, 3)$, $(2, 4), \dots$, $(2, k)$, потом $(3, 4), \dots$, $(3, k)$ и т.д., вплоть до $(k - 1, k)$.

Ядро сильных противоречий можно изобразить графом с вершинами в точках носителя. При этом сильно противоречивые пары задают ребра этого графа. Граф для $S(A, B)$ имеет только одно ребро, оно соединяет вершины 8 и 9, граф содержит одну связную компоненту более чем из одной точки. Граф для $S(A, C)$ имеет 2 ребра, соединяющие вершины 1 и 3, а также 2 и 4, этот граф содержит две связные компоненты более чем из одной точки. В графе для $S(B, C)$ уже 5 ребер, соединяющих вершины 1 и 3, 2 и 3, 2 и 4, 5 и 6, 8 и 9 соответственно. В этом графе три связные компоненты более чем из одной точки, а именно, $\{1, 2, 3, 4\}$ из четырех точек, $\{5, 6\}$ и $\{8, 9\}$ из двух точек каждая.

Каждую кластеризованную ранжировку, как и любое бинарное отношение на носителе из k элементов, можно задать квадратной матрицей $\|x(a, b)\|$ из 0 и 1 порядка k . При этом $x(a, b) = 1$ тогда и только тогда, когда $a < b$ либо $a = b$. В первом случае $x(b, a) = 0$, а во втором $x(b, a) = 1$. При этом хотя бы одно из чисел $x(a, b)$ и $x(b, a)$ равно 1. Из определения сильной противоречивости пары (a, b) вытекает, что для нахождения всех таких пар достаточно поэлементно перемножить матрицы $\|x(a, b)\|$ и $\|y(a, b)\|$, соответствующие двум кластеризованным ранжировкам, и отобрать те и только те пары, для которых $x(a, b)y(a, b) = x(b, a)y(b, a) = 0$, т.е. выделить все пары равных 0 элементов, симметричных относительно главной диагонали.

4. Определение слабой противоречивости

Пусть A и B – две кластеризованные ранжировки. Пару объектов (a, b) назовем *слабо противоречивой* относительно A и B , если упорядоченность этих двух элементов в A и B различается, т.е. 1) $a < b$ в A и $a > b$ или $a = b$ в B , либо 2) $a > b$ в A и $a < b$ или $a = b$ в B , либо 3) $a = b$ в A и $a < b$ или $a > b$ в B . Отметим, что в соответствии с этим определением пара объектов (a, b) гораздо чаще оказывается (слабо) противоречивой, чем в случае сильной противоречивости. Так, эквивалентность $a = b$ в одной кластеризованной ранжировке образует “слабое противоречие” и с $a < b$, и с $a > b$ в другой, и “слабого противоречия” нет тогда и только тогда, когда и во второй кластеризованной ранжировке справедлива эквивалентность $a = b$.

Итак, понятия “сильного” и “слабого” противоречия различаются трактовкой эквивалентности $a = b$. В первом случае она считается совместимой с любым неравенством между объектами и потому исключает противоречие. Следовательно, “сильное противоречие” — это, так сказать, весьма серьезное противоречие, с которым нельзя справиться, нарушая эквивалентность $a = b$ в ту или иную сторону. Во втором случае любое различие мнений относительно рассматриваемых объектов считается противоречием и эквивалентность $a = b$ в одной кластеризованной ранжировке не противоречит лишь такой же эквивалентности во второй.

Совокупность слабо противоречивых пар объектов для двух кластеризованных ранжировок A и B назовем *ядром слабых противоречий* и обозначим $Q(A, B)$. Для рассмотренных выше в качестве примеров трех кластеризованных ранжировок A , B и C имеем

$$\begin{aligned} Q(A, B) &= [\{1, 2\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{4, 5\}, \{5, 6\}, \{5, 7\}, \{6, 7\}, (8, 9), \{8, 10\}], \\ Q(A, C) &= [(1, 3), \{1, 4\}, \{2, 3\}, (2, 4), \{5, 6\}, \{5, 8\}, \{6, 7\}, \{7, 8\}, \{9, 10\}], \\ Q(B, C) &= [\{1, 2\}, (1, 3), (1, 4), (2, 3), (2, 4), \{3, 4\}, \{3, 5\}, \{4, 5\}, \\ &\quad (5, 6), \{5, 7\}, \{5, 8\}, \{7, 8\}, (8, 9), \{8, 10\}, \{9, 10\}]. \end{aligned}$$

Каждое ядро слабых противоречий содержит соответствующее ядро сильных противоречий. Количество пар, находящихся в отношении слабого противоречия, как и следовало ожидать, существенно больше количества пар, находящихся в отношении сильного противоречия.

Ядро слабых противоречий также можно изобразить графом с вершинами в точках носителя. При этом слабо противоречивые пары задают ребра этого графа, которых оказывается много больше, чем в случае изучения сильных противоречий. Граф для $Q(A, B)$ имеет 10 ребер, позволяющих выделить две связанные компоненты, а именно, $\{1, 2, 3, 4, 5, 6, 7\}$ и $\{8, 9, 10\}$. Граф для $Q(A, C)$ состоит из 9 ребер, образующих три связанные компоненты, а именно, $\{1, 2, 3, 4\}$, $\{5, 6, 7, 8\}$ и $\{9, 10\}$. В графе для $Q(B, C)$ имеется 15 ребер, связанная компонента всего одна, она включает все 10 объектов, образующих носитель.

Рассмотрим введенные выше матрицы $\|x(a, b)\|$ и $\|y(a, b)\|$, соответствующие двум кластеризованным ранжировкам. Пара (a, b) не является слабо противоречивой тогда и только тогда, когда $x(a, b) = y(a, b)$ и $x(b, a) = y(b, a)$. Поэтому для выделения слабо противоречивых пар достаточно сложить по модулю 2 матрицы $\|x(a, b)\|$ и $\|y(a, b)\|$. Каждая имеющаяся в сумме $\|x(a, b)\| + \|y(a, b)\|$ единица указывает на слабо противоречивую пару. Напомним, что для получения сильно противоречивых пар матрицы $\|x(a, b)\|$ и $\|y(a, b)\|$ надо поэлементно перемножить, причем каждая пара нулевых элементов, симметричных относительно главной диагонали, указывает на сильно противоречивую пару.

5. Алгоритмы согласования

Предлагаемые алгоритмы согласования некоторого числа кластеризованных ранжировок состоят из трех этапов.

Этап 1. Выделяются противоречивые пары объектов во всех парах кластеризованных ранжировок. Поскольку выше рассмотрены два вида противоречий — сильное и слабое, то алгоритмов согласования также два — сильный и слабый.

Этап 2. Выделяются кластеры итоговой кластеризованной ранжировки (т.е. классы эквивалентности — связанные компоненты графов, соответствующих объединению попарных ядер противоречий).

Этап 3. Эти кластеры (классы эквивалентности) упорядочиваются. Для установления порядка между кластерами произвольно выбирается один объект из первого кластера и второй — из второго, порядок между кластерами устанавливается

такой же, какой имеется между выбранными объектами в любой из рассматриваемых кластеризованных ранжировок (корректность подобного упорядочения, т.е. его независимость от выбора той или иной пары объектов, вытекает из доказанных нами теорем, примеры соответствующей математической техники приведены в Приложении). В случае слабого согласования никакие два объекта из разных кластеров не могут быть равны, поэтому описанная выше процедура не встречает препятствий. В случае сильного согласования это не так – два объекта из разных кластеров *согласованной* кластеризованной ранжировки могут оказаться эквивалентными в одной из исходных кластеризованных ранжировок (т.е. находиться в одном кластере). В таком случае надо рассмотреть упорядоченность этих объектов в какой-либо другой из исходных кластеризованных ранжировок. Если же во всех исходных кластеризованных ранжировках два рассматриваемых объекта находились в одном кластере, то естественно считать (и это является уточнением к этапу 3 алгоритма), что они находятся в одном кластере и в *согласованной* кластеризованной ранжировке.

Результат сильного согласования кластеризованных ранжировок $A, B, C, \dots$ обозначим $f(A, B, C, \dots)$, результат их слабого согласования – $F(A, B, C, \dots)$. На основе проведенных выше рассуждений о ядрах сильных и слабых противоречий нетрудно получить:

$$\begin{aligned} f(A, B) &= [1 < 2 < 3 < 4 < 5 < 6 < 7 < \{8, 9\} < 10], \\ f(A, C) &= [\{1, 3\} < \{2, 4\} < 5 < 6 < 7 < 8 < 9 < 10], \\ f(B, C) &= [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10], \\ f(A, B, C) &= f(B, C) = [\{1, 2, 3, 4\} < \{5, 6\} < 7 < \{8, 9\} < 10], \\ F(A, B) &= [\{1, 2, 3, 4, 5, 6, 7\} < \{8, 9, 10\}], \\ F(A, C) &= [\{1, 2, 3, 4\} < \{5, 6, 7, 8\}, \{9, 10\}], \\ F(B, C) &= [\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}], \\ F(A, B, C) &= F(B, C) = [\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}]. \end{aligned}$$

Полученные результаты показывают весь спектр *согласованных* кластеризованных ранжировок – от $f(A, B)$, в которой только два элемента объединены в кластер, до $F(B, C)$, в которой все элементы составляют один кластер. В случае $f(A, B)$ дополнительного изучения с целью упорядочения требуют только объекты 8 и 9. В случае $F(B, C)$ все объекты объединились в один кластер, т.е. кластеризованные ранжировки оказались настолько противоречивыми, что процедура согласования не позволила провести декомпозицию задачи нахождения итогового мнения экспертов, т.е. разбить объекты на группы, каждую из которых можно анализировать отдельно. Наш опыт анализа реальных кластеризованных ранжировок показал, что почти всегда ситуация ближе к случаю $f(A, B)$, чем к случаю $F(B, C)$.

6. Свойства алгоритмов согласования

6.1. Пусть $D = f(A, B, C, \dots)$. Если $a < b$ в сильно *согласованной* кластеризованной ранжировке D , то $a < b$ или $a = b$ в каждой из исходных кластеризованных ранжировок $A, B, C, \dots$.

6.2. Пусть $D = F(A, B, C, \dots)$. Если $a < b$ в слабо *согласованной* кластеризованной ранжировке D , то $a < b$ в каждой из исходных кластеризованных ранжировок $A, B, C, \dots$.

6.3. Построение *согласованных* кластеризованных ранжировок может осуществляться поэтапно. В частности,

$$f(A, B, C) = f(f(A, B), f(A, C), f(B, C)).$$

Это связано с тем, что ядра сильных и слабых противоречий для нескольких кластеризованных ранжировок являются объединениями таких ядер для всех пар рассматриваемых кластеризованных ранжировок.

6.4. Построение *согласованных* кластеризованных ранжировок нацелено на выделение общего упорядочения в исходных кластеризованных ранжировках. Однако при этом некоторые общие свойства исходных кластеризованных ранжировок могут теряться. Так, при сильном согласовании кластеризованных ранжировок B и C , рассмотренных выше в качестве примеров, противоречия в упорядочении элементов 1 и 2 нет, пара $(1, 2)$ не входит в ядро сильных противоречий $S(B, C)$. В кластеризованной ранжировке B эти объекты входят в один кластер: $1 = 2$, в то время как в кластеризованной ранжировке C имеем: $1 < 2$. Следовательно, при их отдельном рассмотрении можно было принять упорядочение $1 < 2$. Однако в $f(B, C)$ они попали в один кластер, т.е. возможность их упорядочения исчезла. Это связано с поведением объекта 3, который “перескочил” в C на первое место и “увлек с собой” в противоречие”пару $(1, 2)$, образовав противоречивые пары и с 1, и с 2. Другими словами, связная компонента графа, соответствующего ядру противоречий, сама по себе не всегда является полным графом. Недостающие ребра при этом соответствуют парам типа $(1, 2)$, которые сами по себе не являются противоречивыми, но “увлекаются в противоречие” другими парами.

6.5. В *согласованной* ранжировке могут быть кластеры иного происхождения, не связанные с противоречиями. Если во всех исходных кластеризованных ранжировках два элемента не различаются, например, имеется кластер $\{1, 3\}$, то и в *согласованной* ранжировке будет такой кластер, либо он войдет в более обширный кластер. В *согласованной* ранжировке сохраняется не только общая для всех исходных кластеризованных ранжировок упорядоченность, но и общая для всех эквивалентность.

6.6. В наших исследованиях необходимость согласования кластеризованных ранжировок возникла при разработке методики применения экспертных оценок в задачах экологического страхования и химической безопасности биосферы [3–6]. Популярным является метод упорядочения по средним рангам, в котором итоговая ранжировка строится на основе средних арифметических рангов, выставленных отдельными экспертами [1, 2]. Однако из теории измерений известно [11, гл. 3], что более обоснованным является использование не средних арифметических, а медиан (при обобщении порядковых данных типа ответов экспертов применение средних арифметических некорректно, в отличие от медиан). Вместе с тем метод средних рангов весьма известен и широко применяется, так что просто отбросить его нецелесообразно. Поэтому было принято решение об одновременном применении обоих методов. Реализация этого решения потребовала разработки методики согласования двух указанных кластеризованных ранжировок.

6.7. Область применения рассматриваемого в настоящей статье метода не ограничивается экспертными оценками. Он был с успехом применен для сравнения качества математических моделей процесса испарения жидкости. Имелись данные экспериментов и результаты расчетов по 8 математическим моделям. Сравнивать модели можно по различным критериям качества. Например, по сумме модулей относительных отклонений расчетных и экспериментальных значений. Можно в каждой экспериментальной точке упорядочить модели по качеству, а потом получать единую оценку методами средних рангов и медиан. Использовались и иные методы. Затем применялись методы согласования полученных кластеризованных ранжировок. В результате оказалось возможным упорядочить модели по качеству и использовать это упорядочение при разработке банка математических моделей, используемого в задачах химической безопасности биосферы.

7. Заключение

В настоящей работе предложен новый метод анализа экспертных оценок, принципиально отличающийся от ранее известных [1, 2]. Он позволяет проводить декомпозицию задачи экспертного оценивания, разделяя область согласия экспертов и область противоречий в их ответах. Метод позволяет также проводить согласование результатов применения различных алгоритмов обработки ответов экспертов или иных результатов наблюдений. В соответствии с методологией устойчивости [11] выводы, общие для всех использованных методов, могут отражать свойства реальности, в то время как противоречащие друг другу результаты расчетов демонстрируют лишь влияние субъективизма исследователей, выбравших тот или иной метод. Метод используется в учебном процессе Московского государственного технического университета им. Н. Э. Баумана, начиная со второго курса.

ПРИЛОЖЕНИЕ

Некоторые математические результаты

Рассмотрим процедуру слабого согласования. С точки зрения теории бинарных отношений ядро слабых противоречий для двух кластеризованных ранжировок можно назвать квазитолерантностью расхождений (КТР). Предположим, что отсутствуют пары (a, b) такие, что $a = b$ для обеих кластеризованных ранжировок (см. п. 6.5 выше). Каждую такую пару можно условно заменить одним элементом.

Отношение КТР является симметричным (если пара (a, b) входит в него, то входит и пара (b, a)) и антирефлексивным (ни одна пара (a, a) не входит в КТР). Свойством транзитивности это бинарное отношение, вообще говоря, не обладает (если пары (a, b) и (b, c) входят в него, то пара (a, c) может входить в КТР, а может и не входить).

Формально присоединим к КТР все пары вида (a, a) . Получим рефлексивное симметричное отношение, т.е. толерантность [10]. Будем называть ее толерантностью расхождений (ТР).

Построим новое бинарное отношение $\text{Зам}(\text{ТР})$ путем транзитивного замыкания (в смысле теории бинарных отношений [10, с. 27]) толерантности расхождений. Это означает, что подмножество пар объектов, входящих в ТР, пополняется некоторыми новыми парами. А именно, если a, b и c – три объекта такие, что пара (a, b) и пара (b, c) входят в ТР, то пару (a, c) включаем в замыкание этой ТР. Для полученного множества пар повторяем описанную операцию. Продолжаем так до тех пор, пока новые пары не перестанут добавляться (процесс не может продолжаться бесконечно, поскольку общее число пар конечно).

Бинарное отношение $\text{Зам}(\text{ТР})$ можно описать и по-другому: пара (a, b) входит в $\text{Зам}(\text{ТР})$ тогда и только тогда, когда либо она входит в ТР, либо существует конечная последовательность объектов $c, d, e, \dots, q$ такая, что пары (a, c) , (c, d) , $(d, e), \dots, (q, b)$ входят в ТР, т.е. от a к b можно пройти за несколько шагов, каждый из которых – переход от первого элемента пары, входящей в ТР, ко второму.

Последнее замечание подсказывает наглядную геометрическую интерпретацию операции замыкания. Представим себе объекты точками на плоскости. Пара (a, b) входит в ТР тогда и только тогда, когда от a до b можно добраться по дороге. Тогда ясно, что пара (a, c) входит в $\text{Зам}(\text{ТР})$ в том и только в том случае, когда от a до c можно добраться по дороге, возможно через несколько промежуточных пунктов (объектов).

Описание структуры $\text{Зам}(\text{ТР})$ дает следующая теорема.

Теорема о структуре замыкания. Замыкание толерантности расхождений – отношение эквивалентности (рефлексивное симметричное транзитивное отношение), задающее разбиение объектов на кластеры (группы эквивалентных в

рассматриваемом смысле объектов). Кластеры между собой упорядочены: все объекты одного кластера одновременно лучше (или одновременно хуже) всех объектов другого кластера сразу по двум упорядочениям, соответствующим обоим кластеризованным ранжировкам. Внутри же кластеров, состоящих более чем из одного элемента, имеются противоречия: для какого-то объекта есть другой из того же кластера такой, что упорядочение по одной кластеризованной ранжировке противоречит упорядочению по другой кластеризованной ранжировке.

Доказательство. Рефлексивность $\text{Зам}(\text{ТР})$ вытекает из рефлексивности ТР: поскольку любая пара (a, a) входит в ТР, то она входит и в $\text{Зам}(\text{ТР})$. Симметричность вытекает из симметричности ТР: если из a в b можно добраться по цепочке $c, d, e, \dots, q$, то из b в a – по обратной цепочке $q, \dots, e, d, c$, каждые два соседних элемента которой образуют пару, входящую в ТР наряду с “симметричной” парой из прямой цепочки. Транзитивность вытекает из процедуры построения $\text{Зам}(\text{ТР})$. В теории бинарных отношений рефлексивное симметричное и транзитивное отношение называется эквивалентностью [10, с. 54].

Известно [10, с. 55–56, теорема 2.1], что отношение эквивалентности задает такое разбиение множества объектов на кластеры (классы, группы, подмножества), что пара (a, b) входит в $\text{Зам}(\text{ТР})$ тогда и только тогда, когда объекты a и b включены в один и тот же кластер.

Теперь введем упорядоченность кластеров.

Лемма. Пусть $X = \{a, b, \dots\}$ и $Y = \{c, d, \dots\}$ – два кластера. Пусть a меньше c при использовании одной из двух рассматриваемых кластеризованных ранжировок. Тогда a меньше c и при сравнении по второй кластеризованной ранжировке. Более того, любой объект из первого кластера меньше любого объекта из второго кластера в смысле любой из двух кластеризованных ранжировок.

Докажем лемму. Если бы a было больше или равно c по второй кластеризованной ранжировке, то пара (a, c) входила бы в КТР и ТР, а потому объекты a и c входили бы в один класс разбиения, соответствующего $\text{Зам}(\text{ТР})$, что противоречит исходному предположению. Это рассуждение показывает также, что любые два объекта b и d из разных кластеров одинаково упорядочены по двум рассматриваемым кластеризованным ранжировкам.

Однако совпадает ли упорядоченность b и d (или даже b и c) с упорядоченностью a и c ?

Одну из упорядоченностей, задаваемых кластеризованными ранжировками, обозначим знаком $<$ (т.е. “меньше”; знак $>$ означает здесь “больше или равно”). Может ли быть так, что $a < c$, но $b > c$? Тогда $a < c < b$. Вторую упорядоченность обозначим знаком $//$. Тогда, в соответствии с рассуждениями предыдущего абзаца $a//c//b$, следовательно, пара (a, b) не может входить в КТР, а потому и в ТР.

Поскольку a и b лежат в одном кластере, то существует цепочка $a(1) = a, a(2), a(3), \dots, a(t) = b$ такая, что пары $(a(p), a(p+1))$ входят в КТР, $p = 1, 2, 3, \dots, t-1$. Рассмотрим минимальное m такое, что $a(m) < c, a(m+1) > c$ (такое m существует, поскольку $a(1) < c$, в то время как $a(t) > c$). Тогда в рассуждениях предыдущего абзаца можно положить $a = a(m), b = a(m+1)$. Получаем, что пара $(a(m), a(m+1))$ не входит в КТР, что противоречит определению $\text{Зам}(\text{ТР})$.

Итак, доказано, что из $a < c$ вытекает $b < c$ для любого b из кластера, включающего a . Аналогичным образом устанавливается, что $b < d$ для любого d из кластера, включающего c . Лемма доказана.

Каждый из кластеров, порожденных $\text{Зам}(\text{ТР})$, может состоять из одного или нескольких элементов. Внутри кластера из одного элемента противоречий быть не может. Если в кластере несколько элементов, то хотя бы одна пара объектов из этого кластера входит в КТР. Однако некоторые пары могут и не содержать противоречий. Например, если упорядочения имеют вид $a < b < c$ и $c//a//b$, то пары (b, c)

и (a, c) входят в КТР, а пара (a, b) – нет. Если же второе упорядочение имеет вид $c // b // a$, то все три пары входят в КТР.

Теорема о структуре замыкания доказана.

Другие математические утверждения, содержащиеся в основной части статьи, в частности, касающиеся сильного согласования, а также анализа более чем двух кластеризованных ранжировок и др., могут быть доказаны аналогичным образом. Поскольку доказательства не встречаются принципиальных трудностей, но довольно обширны, мы сочли возможным их опустить.

СПИСОК ЛИТЕРАТУРЫ

1. Орлов А. И. Экспертные оценки // Заводская лаборатория. 1996. Т. 62. № 1. С. 54–60.
2. Литвак Б. Г. Экспертные оценки и принятие решений. М.: Патент, 1996.
3. Горский В. Г., Курочкин В. К., Орлов А. И., Швецова-Шиловская Т. Н. О научно-методологическом обеспечении безопасности больших химических систем // Управление большими системами. Материалы Международной научно-практической конференции (22–26 сентября 1997 г., ИПУ РАН, Москва, Россия). М.: СИНТЕГ, 1997. С. 164.
4. Горский В. Г., Орлов А. И., Курочкин В. К., Гриценко А. А. К проблеме классификации сложных опасных систем // Управление большими системами. Материалы Международной научно-практической конференции (22–26 сентября 1997 г., ИПУ РАН, Москва, Россия). М.: СИНТЕГ, 1997. С. 211.
5. Горский В. Г., Курочкин В. К., Моткин Г. А. и др. Методологические основы ранжирования и классификации промышленных объектов, подлежащих экологическому страхованию // Труды Второй Всероссийской конференции “Теория и практика экологического страхования”. М.: Ин-т проблем рынка РАН, 1996. С. 7–12.
6. Орлов А. И., Горский В. Г., Жигарев В. Н. и др. Экспертные оценки: современное состояние и перспективы использования в задачах экологического страхования // Труды Второй Всероссийской конференции “Теория и практика экологического страхования”. М.: Ин-т проблем рынка РАН, 1996. С. 20–23.
7. Холлендер М., Вулф Д. Непараметрические методы статистики. М.: Финансы и статистика, 1983.
8. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения. М.: Сов. радио, 1972.
9. Миркин Б. Г. Проблема группового выбора. М.: Наука, 1974.
10. Шрейдер Ю. А. Равенство, сходство, порядок. М.: Наука, 1971.
11. Орлов А. И. Устойчивость в социально-экономических исследованиях. М.: Наука, 1979.

Статья представлена к публикации членом редколлегии В. А. Лотоцким.

Поступила в редакцию 28.09.98