

О метрических свойствах медианы Кемени*

С. Д. Двоенко, Д. О. Пшеничный
sergedv@yandex.ru; denispshenichny@yandex.ru

Тульский государственный университет, г. Тула

Рассмотрена новая задача построения медианы Кемени с метрическими свойствами. При согласовании экспертных мнений требуется получить ранжирование, наименее отличающееся от остальных и имеющее смысл группового мнения. Медиана Кемени является эквивалентом среднего в шкалах (квази)порядков и свободна от противоречий, связанных с выявлением групповых мнений по правилу большинства (парадокс Эрроу). Известный локально-оптимальный алгоритм построения медианы Кемени основан на вычислении матрицы штрафов. Считая, что ранжирования, представленные парными расстояниями, погружены в евклидово метрическое пространство, можно определить средний элемент как центр такого множества. Такой центральный элемент также является ранжированием и должен иметь такой же смысл, как и медиана Кемени. Разработана процедура формирования скорректированной матрицы штрафов для построения метрической медианы Кемени, совпадающей со средним элементом данного множества.

Ключевые слова: *парные сравнения; метрика; правило большинства; медиана Кемени; ранжирование; парадокс Эрроу*

On metric characteristics of the Kemeny's median*

S. D. Dvoenko and D. O. Pshenichny

Tula State University, Tula

Background: For aggregating of expert's opinions, it is necessary to find the final ranking, which is the least different from others and represents the group opinion. The Kemeny's median appears to be a good idea of the average for scales of (quasi-)orderings and is free of some contradictions concerning the building of a group opinion based on the majority rules (Arrow's paradox). The well-known locally optimal algorithm to find the Kemeny's median depends on pairwise distances between rankings and calculates the so-called loss matrix.

Methods: It is assumed that the rankings represented by pairwise distances between them are immersed as a set in some Euclidean metric space. According to it, one can define the average element as the center of this set. Such central element is a ranking, too, and needs to be similar to the Kemeny's median. To be the Kemeny's median the mathematically correct center of the set of rankings, it needs to be like the center by its distances to other elements. The procedure is developed to build the modified loss matrix and to find the metric Kemeny's median, which coincides with the average element of the given set.

Results: In general, such center element differs by its distances from the corresponding distances of the Kemeny's median to other set elements. The authors find the metric Kemeny's median which coincides with the average element of the given set. Such ranking coincides with the classic Kemeny's median and proves its metric property, or differs from it, if the metric violations in the set configuration appear to be significant.

Concluding Remarks: The metric Kemeny's median is the correct center of the set of rankings and can be used for the correct version of k-means algorithm and others for ordering scales.

*Работа выполнена при финансовой поддержке РФФИ, проект № 15-07-02228.

Keywords: *pairwise comparisons; metrics; majority rule; Kemeny's median; ranking; Arrow's paradox*

1 Введение. Проблема согласования ранжирований

Задача выбора естественным образом заключается в выборе некоторого элемента множества альтернатив, который обладает какими-то наилучшими (с точки зрения авторов) характеристиками.

Хорошо известно, что одно из свойств такой задачи заключается в том, что выбор часто довольно трудно рационально обосновать. Выбор на практике часто основан на интуиции и опыте эксперта. Поэтому и говорят об индивидуальном выборе и предпочтениях индивидуума, т. е. эксперта. Если эксперту удалось выбрать наилучшую альтернативу, то, как правило, ему удастся снова выбрать следующую наилучшую из оставшихся и т. д. В итоге альтернативы оказываются упорядоченными по предпочтениям данного эксперта и образуют ранжирование.

Пусть $A = \{a_1, \dots, a_N\}$ — неупорядоченное множество альтернатив. Если считать, что альтернативы проиндексированы уже после их упорядочения по предпочтениям, то A — упорядоченное множество, представляющее строгое $P = a_1 \succ a_2 \succ \dots \succ a_N$ или в общем случае нестрогое $P = a_1 \succcurlyeq a_2 \succcurlyeq \dots \succcurlyeq a_N$ ранжирование. В последнем случае оказывается, что эксперт может не различать некоторые альтернативы (он не столь категоричен в своих предпочтениях и ставит их на одно место).

Таким образом, получение ранжирования означает, что для всех пар $(a_i, a_j) \in A \times A$ всегда можно указать, какая альтернатива лучше (не хуже) другой. Заметим, что обратное в общем случае неверно. Восстановление ранжирования на основе парных сравнений является отдельной задачей, которая здесь не рассматривается. Таким образом, ранжирование P также может быть представлено матрицей отношений $M_P(N, N)$ с элементами:

$$m_{ij} = \begin{cases} 1, & a_i \succ a_j; \\ 0, & a_i \sim a_j; \\ -1, & a_i \prec a_j. \end{cases}$$

Для двух ранжирований P_u и P_v , представленных своими матрицами отношений M_{P_u} и M_{P_v} , можно вычислить расстояние:

$$d(P_u, P_v) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N |m_{ij}^u - m_{ij}^v| \quad (1)$$

при условии, что элементы из A перечислены одинаково (необязательно в соответствии с одним из этих ранжирований). Известно [1, 2], что эта величина представляет собой метрику на бинарных отношениях линейного (квази)порядка, к которым относятся ранжирования.

Пусть имеется n различных индивидуальных предпочтений (ранжирований). Нужно построить групповое отношение P , согласованное в некотором смысле с отношениями P_1, \dots, P_n . Способы построения группового отношения называют принципами согласования. Существуют разные принципы согласования. Если на способ согласования не наложены ограничения, то он, вообще говоря, может быть любым.

Наиболее распространенным принципом является правило большинства [3]. Например, если P_1, \dots, P_n — ранжирования, то для пары альтернатив, где $a_i \succ a_j$, число

$n(i, j) = \sum_{u=1}^n (m = 1)_{ij}^u$ означает количество таких ранжирований, т. е. количество экспертов, имеющих такое предпочтение. Один из принципов согласования по правилу большинства (мажоритарный) имеет вид: $a_i \succ a_j$ в групповом отношении P , если $n(i, j) \geq n(j, i)$. Если ранжирования строгие, то это эквивалентно условию $n(i, j) \geq n/2$. Известно, что в общем случае мажоритарное отношение может быть нетранзитивным, даже если все индивидуальные предпочтения транзитивны.

Медиана P^* — это такое ранжирование, расстояние от которого до остальных P_1, \dots, P_n является минимальным:

$$P^* = \arg \min_P \sum_{u=1}^n d(P, P_u). \quad (2)$$

Оказывается, если мажоритарное отношение транзитивно (или приведено к такому виду специальными процедурами), то оно является медианой и, в частности, медианой Кемени [1].

2 Построение медианы с метрическими свойствами

Рассмотрим ранжирования P_1, \dots, P_n как некоторое неупорядоченное множество элементов, погруженных некоторым способом в метрическое пространство и представленных матрицей $D(n, n)$ парных расстояний между собой, например, согласно (1). Если нет метрических нарушений в конфигурации элементов множества [4, 5], то центральный элемент P_0 можно представить своими расстояниями до остальных элементов согласно методу Торгерсона:

$$d^2(P_0, P_i) = d_{0i}^2 = \frac{1}{n} \sum_{p=1}^n d_{ip}^2 - \frac{1}{2n^2} \sum_{p=1}^n \sum_{q=1}^n d_{pq}^2, \quad i = 1, \dots, n. \quad (3)$$

Заметим, что метрические нарушения могут сделать невозможным вычисление расстояний (3), представляющих центральный элемент P_0 множества, так как второе слагаемое, представляющее дисперсию множества, может превысить по величине первое слагаемое [4, 5]. В то же время отсутствие второго слагаемого позволит вычислить такие расстояния всегда. Это будут расстояния, представляющие элемент P_{00} , вынесенный за пределы выпуклой оболочки данного множества.

Тогда, согласно известному свойству среднего арифметического, центральный элемент P_0 также оказывается наименее удаленным от всех остальных элементов данного множества и должен формально удовлетворять (2), являясь ранжированием, т. е. медианой.

Проблема заключается в том, что уже построенная медиана P^* представлена как ранжированием, так и своими расстояниями до остальных ранжирований из матрицы $D(n, n)$, добавляя к ней дополнительные строку и столбец. В то же время центральный (средний) элемент P_0 как ранжирование не существует, а представлен только своими расстояниями (3) до остальных элементов множества.

Совпадают ли эти ранжирования $P^* = P_0$? Если это так, то возникает возможность строить математически корректные версии известных алгоритмов распознавания и кластер-анализа для объектов, представленных измерениями в менее мощных (качественных) шкалах, на основе их парных сравнений [6].

В целом, такой подход известен [2, 7, 8]. В рамках такого подхода развивается аксиоматика метрик, введенных на бинарных отношениях, представляющих эквивалентности,

(квази)порядки и т. д. В данном же случае сразу предполагается наличие евклидовой метрики для элементов, погруженных в метрическое пространство.

Известно, что строгие и нестрогие ранжирования формально являются измерениями в шкалах строгого и нестрогого порядка. Корректным преобразованием, переводящим шкалы данного типа друг в друга, является монотонное преобразование, которое, перераспределяя положение альтернатив на числовой оси, не меняет их упорядоченного расположения относительно друг друга.

Достаточно очевидно, что в общем случае элемент P_0 и ранжирование P^* представлены каждый своими расстояниями до остальных элементов P_1, \dots, P_n . Необходимо, используя подходящее монотонное преобразование, показать эквивалентность ранжирований P_0 и P^* . Это можно сделать, показав, что эти расстояния могут быть одинаковыми.

Применим для этого локально-оптимальный алгоритм построения медианы Кемени [1], предложенный в [2].

Известно, что алгоритм построения медианы Кемени основан на вычислении матрицы штрафов $Q(N, N)$, где N — число альтернатив. Пусть некоторое произвольное ранжирование P и ранжирования экспертов P_1, \dots, P_n представлены матрицами отношений M_P и M_{P_1}, \dots, M_{P_n} . Суммарное расстояние от P до всех остальных ранжирований определяется как

$$\sum_{u=1}^n d(P, P_u) = \frac{1}{2} \sum_{u=1}^n \sum_{i=1}^N \sum_{j=1}^N |m_{ij} - m_{ij}^u| = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{u=1}^n |m_{ij} - m_{ij}^u| = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{u=1}^n d_{ij}(P, P_u),$$

где при условии $m_{ij} = 1$ частичные «расстояния» определяются как

$$d_{ij}(P, P_u) = \begin{cases} 0, & m_{ij}^u = 1; \\ 1, & m_{ij}^u = 0; \\ 2, & m_{ij}^u = -1. \end{cases} \quad (4)$$

Элемент матрицы штрафов q_{ij} определяет суммарный штраф на несовпадение предпочтения $a_i \succ a_j$ в неизвестном ранжировании P по сравнению с соответствующим предпочтением в каждом из ранжирований P_1, \dots, P_n :

$$q_{ij} = \sum_{u=1}^n d_{ij}(P, P_u).$$

Алгоритм построения медианы Кемени находит такое упорядочение альтернатив, что сумма элементов матрицы Q над ее диагональю минимальна.

3 Формирование матрицы штрафов

Представляется достаточно очевидным, что частичные «расстояния» (4) достаточно условны. Поэтому также развивается направление, связанное с введением так называемых метризованных бинарных отношений [2], когда элемент $m_{ij} = w_{ij}$ матрицы отношений имеет значения w_{ij} , отличающиеся от $-1, 0, 1$.

Как было сказано выше, для доказательства метричности медианы Кемени необходимо показать, что $P_0 = P^*$, представив P_0 и P^* своими одинаковыми расстояниями до остальных элементов множества. В данном случае появляется возможность применить такое монотонное преобразование шкалы ранжирования для эксперта u , что матрица отношения M_{P_u} , представляющая его ранжирование P_u , естественным образом оказывается

метризованной, так как метрическая медиана (как среднее по множеству) не нарушает метричности конфигурации элементов, представляющих ранжирования экспертов.

Рассмотрим ранжирования P_u, P_0 и P^* , где $\delta = d(P_0, P_u) - d(P^*, P_u) \neq 0$. Пусть $\delta > 0$. Для компенсации такой разницы в расстояниях необходимо равномерно распределить значение $\delta > 0$ по ненулевым элементам матрицы отношений M_{P_u} , сформировав новую матрицу отношений с элементами:

$$m_{ij}^u = \begin{cases} +1 + \frac{2\delta}{k}, & a_i \succ a_j; \\ 0, & a_i \sim a_j; \\ -1 - \frac{2\delta}{k}, & a_i \prec a_j, \end{cases}$$

где $k = N^2 - N - N_0$ — общее число ненулевых элементов без главной диагонали, N_0 — число нулевых недиагональных элементов $m_{ij}^u = 0$. Очевидно, что такая матрица отношений не изменяет ранжирования эксперта P_u .

В этом случае при $m_{ij} = 1$ предположение $a_i \succ a_j$ в неизвестном ранжировании P штрафует эксперта с формированием частичных расстояний:

$$d_{ij}(P, P_u) = \begin{cases} \frac{2\delta}{k}, & m_{ij}^u = +1 + \frac{2\delta}{k}; \\ 1, & m_{ij}^u = 0; \\ 2 + \frac{2\delta}{k}, & m_{ij}^u = -1 - \frac{2\delta}{k}. \end{cases}$$

Пусть $\delta < 0$. Для компенсации такой разницы в расстояниях также необходимо равномерно распределить значение $\delta < 0$ по ненулевым элементам матрицы отношений M_{P_u} . Кроме того, множество изменяемых элементов $k' = k - \Delta k$ дополнительно уменьшается за счет Δk совпадающих элементов $m_{ij}^u = m_{ij}^*$ в матрицах отношений, представляющих ранжирование эксперта P_u и медиану P^* . Действительно, в этом случае $d_{ij}(P^*, P_u) = 0$ и уменьшить его невозможно.

Добавление любой отрицательной величины к значению m_{ij}^u в этом случае будет означать изменение ранжирования эксперта, при котором расстояние между ранжированиями P_u и P^* только возрастет.

Если величина $-1 < 2\delta/k' < 0$, то можно, не изменяя ранжирования P_u , сформировать новую матрицу отношений M_{P_u} с элементами:

$$m_{ij}^u = \begin{cases} +1 - \left| \frac{2\delta}{k'} \right|, & a_i \succ a_j; \\ 0, & a_i \sim a_j; \\ -1 + \left| \frac{2\delta}{k'} \right|, & a_i \prec a_j; \\ m_{ij}^*, & m_{ij}^u = m_{ij}^*, \end{cases}$$

где при $m_{ij} = 1$ предположение $a_i \succ a_j$ в неизвестном ранжировании P штрафует эксперта с формированием частичных расстояний:

$$d_{ij}(P, P_u) = \begin{cases} 0, & m_{ij}^u = 1; \\ 1, & m_{ij}^u = 0; \\ 2, & m_{ij}^u = -1; \\ \left| \frac{2\delta}{k'} \right|, & m_{ij}^u = +1 - \left| \frac{2\delta}{k'} \right|; \\ 2 - \left| \frac{2\delta}{k'} \right|, & m_{ij}^u = -1 + \left| \frac{2\delta}{k'} \right|. \end{cases}$$

Если величина $2\delta/k' \leq -1$, то корректировка элементов матрицы отношений M_{P_u} неизбежно изменит знак некоторых элементов m_{ij}^u , что означает изменение ранжирования P_u . Так как изменять экспертное ранжирование мы не имеем права, то следует изменить ранжирование, соответствующее медиане P^* . Фактически это означает, что ранжирования P_0 и P^* различны. Поэтому, изменяя P^* , найдем другое ранжирование, лучше соответствующее P_0 . В этом случае при $-2 \leq 2\delta/k' \leq -1$ формируется новая матрица отношений M_{P^*} с элементами:

$$m_{ij}^* = \begin{cases} +1 - \left| \frac{2\delta}{k'} \right|, & a_i \succ a_j; \\ 0, & a_i \sim a_j; \\ -1 + \left| \frac{2\delta}{k'} \right|, & a_i \prec a_j; \\ m_{ij}^u, & m_{ij}^u = m_{ij}^*, \end{cases}$$

где при $m_{ij} = 1$ предположение $a_i \succ a_j$ в неизвестном ранжировании P штрафуетя ранжированием P^* с формированием частичных расстояний:

$$d_{ij}(P, P^*) = \begin{cases} 0, & m_{ij}^* = 1; \\ 1, & m_{ij}^* = 0; \\ 2, & m_{ij}^* = -1; \\ \left| \frac{2\delta}{k'} \right|, & m_{ij}^* = +1 - \left| \frac{2\delta}{k'} \right|; \\ 2 - \left| \frac{2\delta}{k'} \right|, & m_{ij}^* = -1 + \left| \frac{2\delta}{k'} \right|. \end{cases}$$

Если величина $2\delta/k' < -2$, то распределим по ранее указанным k' элементам только величину -2 . Тогда отрицательный остаток $(2\delta/k' + 2)k' = 2\delta + 2k' < 0$ следует распределить по k_0 нулевым элементам $m_{ij}^* = 0$ матрицы отношений для ранжирования P^* . В итоге формируется новая матрица отношений M_{P^*} с элементами:

$$m_{ij}^* = \begin{cases} +1 - 2, & a_i \succ a_j; \\ -\left| \frac{2\delta}{k_0} + \frac{2k'}{k_0} \right|, & a_i \sim a_j; \\ -1 + 2, & a_i \prec a_j; \\ m_{ij}^u, & m_{ij}^u = m_{ij}^*, \end{cases}$$

где при $m_{ij} = 1$ предположение $a_i \succ a_j$ в неизвестном ранжировании P штрафуетя ранжированием P^* с формированием частичных расстояний:

$$d_{ij}(P, P^*) = \begin{cases} 0, & m_{ij}^* = 1; \\ 1, & m_{ij}^* = 0; \\ 2, & m_{ij}^* = -1; \\ 1 + \left| \frac{2\delta}{k_0} + \frac{2k'}{k_0} \right|, & m_{ij}^* = -\left| \frac{2\delta}{k_0} + \frac{2k'}{k_0} \right|. \end{cases}$$

Легко увидеть, что после такой корректировки матрицы отношений M_{P^*} новое ранжирование P^* , вообще говоря, уже не является медианой, так как в нем произошли перестановки (инверсии) на некоторых парах альтернатив.

При вычислении матрицы штрафов Q учитываются только измененные матрицы отношений для ранжирований P_1, \dots, P_n и измененные матрицы для ранжирования P^* в тех случаях, когда оно корректировалось вместо соответствующего ранжирования эксперта.

4 Эксперименты

4.1 Исходные данные о ранжировании проектов

В качестве экспериментальных данных были рассмотрены материалы исследования 2000 г. по оценке и выбору 14 проектов, обеспечивающих достижение стратегических целей ОАО «Газпром» [9]. В табл. 1 приведен перечень инвестиционных проектов, в осуществлении которых предполагалось участие «Газпрома». Указанный перечень был сформирован по материалам открытой и зарубежной печати.

Таблица 1 Перечень инвестиционных проектов

№	Проекты	Содержание проектов
1	Южный парс	Освоение двух новых нефтегазоносных участков на месторождении «Южный Парс» (Иран)
2	Голубой поток	Строительство морского участка газопровода (двух ниток протяженностью 380 км) и компрессорной станции «Береговая»
3	Ямал–Европа	Завершение строительства трансконтинентального трубопровода для транспортировки сибирского газа в Европу
4	Псковская ГРЭС	Приобретение в счет долгов «Газпрому» Псковской ГРЭС и завершение строительства 3-го энергоблока
5	Метан Кузбасс	Создание компанией «Метан Кузбасса» и администрацией Кемеровской области опытно-промышленного производства по добыче метана из угольных пластов
6	Приразломное	Освоение совместно с партнерами арктического нефтяного месторождения «Приразломное»
7	Трансбалканский трубопровод	Строительство газокompрессорной станции на Трансбалканском газопроводе. Расширение мощностей транспортировки газа на Трансбалканском газопроводе
8	Газопровод Петрозаводск–Кондопога	Строительство в Карелии 68-километрового трубопровода для обеспечения газом Кондопожского целлюлозно-бумажного комбината и жителей севера республики
9	Экология	Совместный проект сокращения эмиссии углекислого газа на участке «Ужгородского коридора» (Волготрансгаз)
10	Дегазификация энергетики	Снижение расхода газа на нужды электроэнергетики
11	Штокмановское	Освоение Штокмановского газоконденсатного месторождения, разведка и освоение добычи газа и газового конденсата Штокмановского газоконденсатного месторождения
12	Газификация автотранспорта	Газификация автотранспорта в Новосибирской области, отказ от бензина, использование пропан-бутановой смеси метана
13	Космическая связь	Создание группировки спутниковой связи для обеспечения «Газпрома» коммерческой технологической связью
14	АСУ корпоративными финансами	Создание для ОАО «Газпром» автоматизированной системы управления корпоративными финансами

Таблица 2 Стандартизированные ранги инвестиционных проектов

№	Проекты	1	2	3	4	5	6	7	8	МК
1	Южный парс	5	4	14	8	11	12	4,5	5,5	9
2	Голубой поток	1	1	3,5	4	12,5	14	9,5	10	4
3	Ямал-Европа	3	3	8,5	3	14	3	11	7	10
4	Псковская ГРЭС	9	7	10	8	9	4,5	4,5	2	2
5	Метан Кузбасс	8	9,5	5,5	8	2,5	7,5	4,5	11	5
6	Приразломное	6	6	8,5	8	9	9	12,5	13	13
7	Трансбалканский трубопровод	10	5	12,5	13,5	6	4,5	4,5	4	7
8	Газопровод Петрозаводск–Кондопога	11	12,5	11	1	2,5	2	4,5	8,5	3
9	Экология	13	12,5	5,5	12	2,5	2	4,5	3	6
10	Дегазификации энергетики	12	12,5	1	8	12,5	6	14	1	14
11	Штокмановское	2	2	2	2	6	7,5	12,5	14	1
12	Газификация автотранспорта	4	8	12,5	8	9	10,5	4,5	5,5	8
13	Космическая связь	7	9,5	7	8	6	10,5	4,5	12	11
14	АСУ корпоративными финансами	14	12,5	3,5	13,5	2,5	2	9,5	8,5	12

Проекты ранжировались по 8 критериям двух видов (выгоды и негативных эффектов):

- 1) финансово-экономическая выгода;
- 2) выгода от изменения конъюнктуры рынка;
- 3) производственно-технологическая выгода;
- 4) социально-политическая выгода;
- 5) политические негативные последствия;
- 6) негативные последствия рыночной конкуренции;
- 7) социальные негативные последствия;
- 8) негативные финансово-экономические последствия.

По характеристикам 1–4 один проект является предпочтительнее другого, если соответствующее значение критерия имеет значение больше, чем у другого проекта. По характеристикам 5–8 проект тем предпочтительнее, чем меньше значение имеет соответствующий критерий. Наиболее предпочтительная альтернатива получает ранг 1, следующая за ней — 2 и т. д. Возможны случаи, когда две и более альтернатив имеют одинаковые значения критериев, т. е. одинаковый ранг. Нам удобно рассматривать критерии как экспертов, а ранжирования — как результат экспертизы. В табл. 2 представлены 8 экспертных ранжирований 14 проектов, где ранжирования представлены стандартизированными рангами, так как некоторые проекты имели одинаковый ранг.

Алгоритм построения медианы Кемени для матрицы штрафов:

$$\begin{pmatrix} 0 & 8 & 8 & 10 & 8 & 7 & 9 & 9 & 9 & 7 & 12 & 11 & 8 & 6 \\ 8 & 0 & 6 & 8 & 6 & 4 & 8 & 10 & 8 & 7 & 8 & 8 & 6 & 8 \\ 8 & 10 & 0 & 8 & 8 & 5 & 8 & 8 & 10 & 8 & 12 & 8 & 8 & 8 \\ 6 & 8 & 8 & 0 & 8 & 8 & 6 & 7 & 7 & 5 & 10 & 5 & 8 & 6 \\ 8 & 10 & 8 & 8 & 0 & 5 & 7 & 8 & 7 & 7 & 9 & 8 & 5 & 7 \\ 9 & 12 & 11 & 8 & 11 & 0 & 10 & 10 & 10 & 7 & 13 & 8 & 9 & 10 \\ 7 & 8 & 8 & 10 & 9 & 6 & 0 & 9 & 11 & 6 & 9 & 6 & 8 & 7 \\ 7 & 6 & 8 & 9 & 8 & 6 & 7 & 0 & 8 & 5 & 6 & 7 & 7 & 6 \\ 7 & 8 & 6 & 9 & 9 & 6 & 5 & 8 & 0 & 9 & 8 & 7 & 7 & 5 \\ 9 & 9 & 8 & 11 & 9 & 9 & 10 & 11 & 7 & 0 & 10 & 9 & 9 & 7 \\ 4 & 8 & 4 & 6 & 7 & 3 & 7 & 10 & 8 & 6 & 0 & 4 & 5 & 8 \\ 5 & 8 & 8 & 11 & 8 & 8 & 10 & 9 & 9 & 7 & 12 & 0 & 7 & 6 \\ 8 & 10 & 8 & 8 & 11 & 7 & 8 & 9 & 9 & 7 & 11 & 9 & 0 & 8 \\ 10 & 8 & 8 & 10 & 9 & 6 & 9 & 10 & 11 & 9 & 8 & 10 & 8 & 0 \end{pmatrix}$$

дает ранжирование P^* (МК), которое также представлено в табл. 2.

4.2 Построение ранжирования, представленного средним объектом

Рассмотрим теперь матрицу парных расстояний между ранжированиями:

$$\begin{pmatrix} 0 & 27 & 98 & 50 & 131 & 156 & 102 & 124 \\ 27 & 0 & 105 & 65 & 128 & 137 & 101 & 113 \\ 98 & 105 & 0 & 88 & 83 & 82 & 128 & 106 \\ 50 & 65 & 88 & 0 & 109 & 116 & 96 & 112 \\ 131 & 128 & 83 & 109 & 0 & 37 & 61 & 101 \\ 156 & 137 & 82 & 116 & 37 & 0 & 80 & 74 \\ 102 & 101 & 128 & 96 & 61 & 80 & 0 & 74 \\ 124 & 113 & 106 & 112 & 101 & 74 & 74 & 0 \end{pmatrix}$$

и проверим метричность конфигурации множества ранжирований как множества элементов, погруженных в метрическое пространство.

Скорректируем метрические нарушения согласно технологии, разработанной в [4, 5]. Определим элемент, вынесенный за пределы выпуклой оболочки данного множества, и представим его своими расстояниями до остальных элементов:

$$(100,07 \ 96,243 \ 93,278 \ 87,725 \ 91,928 \ 97,807 \ 87,893 \ 95,576).$$

Построим матрицу нормированных скалярных произведений относительно данного элемента как начала координат и определим ее собственные числа:

$$(-0,25079 \ -0,11443 \ 0,11728 \ 0,23633 \ 0,57857 \ 1,0547 \ 2,3446 \ 4,0338).$$

Так как среди них есть отрицательные, то скорректируем эту матрицу скалярных произведений, получив только положительные собственные числа:

$$(0,01098 \ 0,04016 \ 0,11631 \ 0,19909 \ 0,47382 \ 1,03831 \ 2,18110 \ 3,94022).$$

Восстановим скорректированную матрицу расстояний:

$$\begin{pmatrix} 0 & 41,281 & 98 & 50 & 125,72 & 143,32 & 102 & 124 \\ 41,281 & 0 & 103,35 & 64,011 & 127,52 & 137,96 & 99,346 & 113,92 \\ 98 & 103,35 & 0 & 88 & 95,464 & 97,307 & 128 & 106 \\ 50 & 64,011 & 88 & 0 & 108,47 & 120,33 & 96 & 112 \\ 125,72 & 127,52 & 95,464 & 108,47 & 0 & 53,025 & 81,431 & 90,68 \\ 143,32 & 137,96 & 97,307 & 120,33 & 53,025 & 0 & 94,451 & 73,325 \\ 102 & 99,346 & 128 & 96 & 81,431 & 94,451 & 0 & 74 \\ 124 & 113,92 & 106 & 112 & 90,68 & 73,325 & 74 & 0 \end{pmatrix}.$$

Так как медиана Кемени построена (МК, см. табл. 2), то определим расстояния от нее до остальных ранжирований по их матрицам отношений. Для скорректированной матрицы расстояний вычислим средний элемент P_0 (СР) и представим его своими расстояниями до остальных элементов множества. Также вычислим разницу δ в расстояниях от среднего элемента и медианы Кемени до остальных элементов множества (табл. 3).

Как было сказано выше, нужно добиться, чтобы не было разницы в расстояниях от медианы Кемени и от среднего элемента до остальных элементов множества. Так как расстояние между ранжированиями определяется по их матрицам отношений, то необходимо распределить соответствующую δ равномерно по элементам матрицы отношений эксперта.

Равномерная корректировка матрицы отношений для ранжирования каждого эксперта выполняется соответствующим удвоенным средним значением $2\delta/k$ (см. табл. 3), так как

Таблица 3 Корректировка матриц отношений для ранжирований экспертов

МК	СР	Разница, δ	Число корректируемых элементов, k	Среднее, $2\delta/k$
76	70,625	-5,375	76	-0,14144
73	69,458	-3,542	66	-0,10732
92	69,102	-22,898	88	-0,52041
66	57,353	-8,647	44	-0,39302
71	65,747	-5,253	58	-0,18113
84	74,666	-9,334	78	-0,23934
68	62,120	-5,880	38	-0,3095
88	66,313	-21,687	86	-0,50435

для вычисления расстояния между двумя ранжированиями каждое различие между парой элементов в матрице отношений учитывается дважды.

Как уже было отмечено, это можно сделать, если такое распределение не изменяет знаков элементов матрицы отношений для соответствующего ранжирования. В этом случае мы не меняем ранжирования эксперта, так как, вообще говоря, мы и не имеем права этого делать. Поэтому для равномерной корректировки выбираются только ненулевые элементы матрицы отношений, так как нуль означает, что два соответствующих проекта занимают одно и то же место в ранжировании эксперта. Такая корректировка естественным образом применяется для компенсации положительной разницы в расстоянии между ранжированиями, так как оно увеличивается.

При распределении отрицательной разницы (см. табл. 3) число корректируемых элементов матрицы отношений дополнительно уменьшается, так как корректируются ненулевые элементы и дополнительно только те, которые отличаются от соответствующих элементов в матрице отношений для медианы. Очевидно, что только в этом случае различие можно уменьшить, учитывая отрицательную разницу. Если при этом знаки элементов матрицы отношений для ранжирования эксперта не изменяются, то его ранжирование не изменяется. Если изменяются, то нельзя можем корректировать матрицу отношений для ранжирования эксперта, так как это означает изменение его ранжирования. Поэтому будем корректировать соответствующие элементы матрицы отношений для медианы. Действительно, это ранжирование все равно должно измениться. В этом случае, если величина $-2 \leq 2\delta/k < 0$, то эта разница распределяется по ненулевым элементам и отличающимся от соответствующих элементов матрицы отношений для ранжирования эксперта. Если величина $2\delta/k < -2$, то распределим по таким элементам матрицы отношений для медианы только величину -2 .

Естественно, что вся отрицательная разница еще не будет скомпенсирована, так как останется еще распределить остаток отрицательной разницы $(2\delta/k + 2)k$ по k_0 элементам, где k_0 — это число нулевых элементов в матрице отношений. Можно показать, что при $k_0 = 0$ величина $2\delta/k$ распределится полностью, а при $k_0 \neq 0$ остаток распределится полностью.

При формировании матрицы штрафов для алгоритма построения медианы Кемени будем, как и раньше, штрафовать наше предположение о предпочтительности одного проекта перед другим по всем возможным парам, когда соответствующий элемент неизвестной матрицы отношений равен 1. Если при этом корректировалась матрица отношений для ранжирования эксперта, то его ранжирование штрафует наши предположения. Если кор-

Таблица 4 Корректировка матриц отношений для ранжирований экспертов

МК	ВП	Разница, δ	Число корректируемых элементов, k	Среднее, $2\delta/k$
76	97,426	21,426	182	0,23545
73	96,583	23,583	168	0,28075
92	96,327	4,327	174	0,049739
66	88,28	22,28	138	0,32289
71	93,95	22,95	156	0,29423
84	100,39	16,393	170	0,19286
68	91,448	23,448	122	0,38439
88	94,346	6,346	178	0,071308

ректировалась матрица отношений для медианы, то только это ранжирование штрафует наши предположения.

В итоге, алгоритм построения медианы Кемени для измененной матрицы штрафов

$$\left(\begin{array}{cccccccccccccccc} 0 & 9,23 & 6,84 & 10,25 & 8,75 & 6,06 & 9,64 & 9,75 & 9,64 & 5,74 & 12,81 & 11,11 & 7,06 & 5,06 \\ 6,77 & 0 & 4,86 & 9,162 & 5,27 & 3,58 & 6,77 & 10,77 & 6,77 & 5,74 & 9,06 & 6,77 & 5,27 & 7,08 \\ 9,16 & 11,14 & 0 & 9,16 & 9,15 & 4,58 & 9,16 & 9,27 & 10,64 & 6,55 & 12,81 & 9,16 & 6,75 & 6,75 \\ 5,75 & 6,84 & 6,84 & 0 & 7,16 & 7,23 & 5,71 & 6,19 & 6,06 & 3,98 & 11,05 & 4,86 & 7,16 & 5,06 \\ 7,25 & 10,73 & 6,85 & 8,84 & 0 & 4,75 & 6,15 & 8,77 & 6,26 & 5,74 & 9,995 & 7,25 & 4,86 & 5,74 \\ 9,94 & 12,42 & 11,42 & 8,77 & 11,25 & 0 & 11,06 & 10,77 & 10,64 & 5,74 & 13,50 & 8,87 & 9,49 & 10,64 \\ 6,36 & 9,23 & 6,84 & 10,29 & 9,85 & 4,95 & 0 & 9,75 & 11,25 & 4,58 & 10,05 & 5,47 & 6,95 & 6,06 \\ 6,25 & 5,23 & 6,73 & 9,81 & 7,23 & 5,23 & 6,25 & 0 & 6,98 & 3,98 & 7,63 & 6,25 & 6,23 & 5,48 \\ 6,36 & 9,23 & 5,36 & 9,94 & 9,74 & 5,36 & 4,75 & 9,02 & 0 & 7,44 & 9,23 & 6,36 & 6,36 & 4,48 \\ 10,26 & 10,26 & 9,45 & 12,03 & 10,26 & 10,26 & 11,42 & 12,03 & 8,56 & 0 & 11,26 & 10,26 & 10,26 & 8,56 \\ 3,19 & 6,94 & 3,19 & 4,95 & 6,01 & 2,50 & 5,95 & 8,37 & 6,77 & 4,74 & 0 & 3,19 & 4,19 & 6,77 \\ 4,89 & 9,23 & 6,84 & 11,14 & 8,75 & 7,13 & 10,53 & 9,75 & 9,64 & 5,74 & 12,81 & 0 & 6,30 & 5,06 \\ 8,94 & 10,73 & 9,25 & 8,84 & 11,14 & 6,51 & 9,05 & 9,77 & 9,64 & 5,74 & 11,81 & 9,70 & 0 & 6,55 \\ 10,94 & 8,92 & 9,25 & 10,94 & 10,26 & 5,36 & 9,94 & 10,52 & 11,52 & 7,44 & 9,23 & 10,94 & 9,45 & 0 \end{array} \right)$$

дает ту же медиану Кемени (см. табл. 2), у которой ее расстояния до остальных ранжирований не отличаются от расстояний среднего элемента до остальных элементов множества из табл. 3.

Таким образом, мы показали, что среднее множества элементов, представляющих ранжирования, также представленное ранжированием, является медианой Кемени.

4.3 Построение ранжирования, представленного произвольным объектом

Снова построим медиану Кемени для ранжирований экспертов и определим ее расстояния до остальных ранжирований. Снова скорректируем исходную матрицу расстояний между ранжированиями, устранив метрические нарушения в конфигурации элементов множества.

Рассмотрим теперь элемент P_{00} (ВП), вынесенный за пределы выпуклой оболочки множества элементов, представляющих ранжирования экспертов. Представим его своими расстояниями до остальных элементов множества (табл. 4). Найдем соответствующее

Таблица 5 Построенные ранжирования

№	ВП	СР	МК
1	11	9	9
2	5	4	4
3	10	10	10
4	2	2	2
5	6	5	5
6	14	13	13
7	7	7	7
8	3	3	3
9	4	6	6
10	12	14	14
11	1	1	1
12	8	8	8
13	9	11	11
14	13	12	12

ему ранжирование, также применив алгоритм построения медианы Кемени к соответствующим образом измененной матрице штрафов.

Алгоритм построения медианы Кемени для измененной матрицы штрафов:

$$\begin{pmatrix} 0 & 6,17 & 7,95 & 8,88 & 6,88 & 6,56 & 7,55 & 7,55 & 7,55 & 6,12 & 10,17 & 9,95 & 7,95 & 5,24 \\ 8,05 & 0 & 5,34 & 6,17 & 5,91 & 3,14 & 8,05 & 8,17 & 8,05 & 6,09 & 6,17 & 8,05 & 5,91 & 7,72 \\ 6,17 & 8,17 & 0 & 6,17 & 6,17 & 4,19 & 6,17 & 6,17 & 8,17 & 7,38 & 10,17 & 6,17 & 8,01 & 8,01 \\ 5,91 & 7,95 & 7,95 & 0 & 8,03 & 7,92 & 5,90 & 7,17 & 6,63 & 3,73 & 8,17 & 4,64 & 8,03 & 5,24 \\ 8,05 & 8,17 & 7,99 & 6,88 & 0 & 4,52 & 6,64 & 6,85 & 6,43 & 6,12 & 7,36 & 8,05 & 4,63 & 6,09 \\ 7,49 & 10,17 & 9,22 & 6,79 & 9,49 & 0 & 8,17 & 8,17 & 8,17 & 6,12 & 11,55 & 6,79 & 7,49 & 8,17 \\ 7,23 & 6,17 & 7,95 & 8,75 & 7,55 & 5,38 & 0 & 7,55 & 9,55 & 5,06 & 7,462 & 5,72 & 8,06 & 6,56 \\ 6,73 & 5,30 & 7,44 & 7,55 & 7,99 & 5,30 & 6,73 & 0 & 7,56 & 3,69 & 4,17 & 6,73 & 6,68 & 5,11 \\ 7,23 & 6,17 & 5,85 & 7,55 & 7,90 & 5,85 & 4,59 & 7,32 & 0 & 8,81 & 6,17 & 7,23 & 7,23 & 4,04 \\ 7,49 & 7,46 & 6,17 & 9,49 & 7,49 & 7,49 & 8,17 & 9,45 & 5,45 & 0 & 8,17 & 7,49 & 7,49 & 5,45 \\ 3,08 & 8,11 & 3,08 & 5,47 & 6,86 & 1,70 & 6,76 & 10,7 & 8,05 & 4,80 & 0 & 3,08 & 4,37 & 8,05 \\ 4,51 & 6,17 & 7,95 & 10,17 & 6,88 & 7,83 & 8,60 & 7,55 & 7,55 & 6,12 & 10,17 & 0 & 6,76 & 5,24 \\ 6,88 & 8,17 & 6,17 & 6,88 & 10,16 & 6,91 & 6,85 & 7,55 & 7,55 & 6,12 & 9,46 & 8,07 & 0 & 7,38 \\ 8,17 & 6,60 & 6,17 & 8,17 & 7,46 & 5,85 & 7,49 & 9,01 & 9,94 & 8,81 & 6,17 & 8,17 & 6,17 & 0 \end{pmatrix}$$

дает новое ранжирование (ВП), показанное в табл. 5. Легко увидеть, что данное ранжирование отличается от медианы Кемени, расстояние между данным ранжированием и медианой Кемени по матрицам их отношений равно 14.

5 Заключение

В экспериментах было показано, что одному и тому же ранжированию могут соответствовать разные векторы расстояний до остальных элементов данного множества (ранжирований), именно поэтому авторам удалось доказать метричность медианы Кемени на представленных данных. Очевидно, что такое свойство обеспечивается монотонным преобразованием ранговой шкалы, не нарушающим порядок элементов множества.

Также было показано, что вектор расстояний элемента, сильно отличающегося от среднего (центрального) элемента множества, формирует другое ранжирование, не совпадающее с ранжированием для среднего объекта.

В общем случае можно показать, что, и наоборот, разным ранжированиям могут соответствовать одинаковые векторы расстояний до других элементов множества (ранжирований). В частности, используя алгоритм построения медианы Кемени, можно в каждом случае построить ранжирование, имеющее такие же расстояния до остальных, как и индивидуальное ранжирование эксперта. В этом случае ранжирование, которое является медианой относительно исходной матрицы штрафов, неизбежно подвергнется немонотонному преобразованию, которое изменит упорядочение элементов в данном ранжировании.

Литература

- [1] Кемени Дж., Снелл Дж. Кибернетическое моделирование. — М.: Сов. Радио, 1972. 192 с.
- [2] Литвак Б. Г. Экспертная информация: методы получения и анализа. — М.: Радио и связь, 1982. 184 с.
- [3] Миркин Б. Г. Проблема группового выбора. — М.: Наука, 1974. 256 с.
- [4] Двоенко С. Д., Пшеничный Д. О. О метрической коррекции матриц парных сравнений // Машинное обучение и анализ данных, 2013. Т. 1. № 5. С. 606–620. <http://jmllda.org/papers/doc/2013/no5/>.
- [5] Двоенко С. Д., Пшеничный Д. О. Оптимальная коррекция метрических нарушений в матрицах парных сравнений // Машинное обучение и анализ данных, 2014. Т. 1. № 7. С. 885–890. <http://jmllda.org/papers/doc/2014/no7/>.
- [6] Двоенко С. Д. Кластеризация множества, описанного парными расстояниями и близостями между его элементами // Сибирский журнал индустриальной математики, 2009. Т. 12. № 1(37). С. 61–73.
- [7] Миркин Б. Г. Анализ качественных признаков. — М.: Статистика, 1976. 166 с.
- [8] Миркин Б. Г. Анализ качественных признаков и структур. — М.: Статистика, 1980. 319 с.
- [9] http://www.bastion.ru/files/materials/analit/Ekonom/Komplex_ocenka_invest_Gazprom.doc.

References

- [1] Kemeny, J., and J. Snell. 1963. *Mathematical models in the social sciences*. New York, NY: Blaisdell. 145 p.
- [2] Litvak, B. G., 1982. *Expert information: Methods of acquisition and analysis*. Moscow: Radio i svyaz'. 184 p. (In Russian.)
- [3] Mirkin, B. G. 1974. *The problem of group choice*. Moscow: Nauka. 256 p. (In Russian.)
- [4] Dvoenko, S. D., and D. O. Pshenichny., 2013. On metric correction of matrices of pairwise comparisons. *JMLDA* 1(5):606–620. (In Russian.) Available at: <http://jmllda.org/papers/doc/2013/no5/> (accessed November 3, 2015).
- [5] Dvoenko, S. D., and D. O. Pshenichny. 2014. Optimal correction of metrical violations in matrices of pairwise comparisons. *JMLDA* 1(7):885–890. (In Russian.) Available at: <http://jmllda.org/papers/doc/2014/no7/> (accessed November 3, 2015).
- [6] Dvoenko, S. D., 2009. Clustering and separating of a set of members in terms of mutual distances and similarities. *Trans. Machine Learning Data Mining* 2(2):80–99.
- [7] Mirkin, B. G. 1976. Analysis of qualitative attributes. Moscow: Statistics. 166 p. (In Russian.)
- [8] Mirkin, B. G. 1980. Analysis of qualitative attributes and structures. Moscow: Statistics. 319 p. (In Russian.)
- [9] http://www.bastion.ru/files/materials/analit/Ekonom/Komplex_ocenka_invest_Gazprom.doc.