О метрической коррекции матриц парных сравнений*

Двоенко С. Д.[1]**, *Пшеничный Д. О.***[2]
[1]dsd@tsu.tula.ru, [2]denispshenichny@yandex.com
[1,2]Тульский государственный университет

В задачах интеллектуального анализа экспериментальные данные часто сразу представлены результатами парных сравнений объектов между собой. В отсутствие исходного признакового пространства условием корректного погружения данного множества объектов в метрическое пространство является неотрицательная определенность матрицы парных близостей элементов множества друг к другу. В этом случае близости интерпретируются как скалярные произведения, а соответствующие различия — как расстояния. В работе рассмотрены условия возникновения метрических нарушений и предложен подход к коррекции метрических нарушений в матрицах парных сравнений за счет минимальных изменений значений некоторых их элементов.

Ключевые слова: *метрика, расстояние, близость, собственные числа, детерминант, парные сравнения.*

On metric correction of matrices of pairwise comparisons*

Dvoenko S. D.[1]**, *Pshenichny D. O.***[2]
[1,2]State University of Tula

In machine learning and data mining, the experimental results are often immediately represented as matrices of pairwise comparisons between set elements. The condition of the correct immersion of the given set of objects into a metric space is a nonnegative definiteness of the pairwise similarity matrix. In this case, similarities are interpreted as scalar products and dissimilarities as distances, respectively. In this paper, the metric violation conditions are under investigation. The approach to correct metric violations in pairwise comparison matrices is developed based on an idea to minimize distortions of some elements.

Keywords: *metrics, distance, similarity, eigenvalues, determinant, pairwise comparisons.*

Введение

Условием корректного погружения множества в метрическое пространство является неотрицательная определенность матрицы парных близостей его элементов [1]. В этом случае, применяя соответствующие «беспризнаковые» версии алгоритмов машинного обучения и кластер-анализа [2, 3] к матрице сходства (скалярные произведения или неотрицательные величины — близости) или к соответствующей ей матрице различий (расстояния), мы получим математически корректный результат обработки.

В данной работе рассматривается подход к регулируемой коррекции матриц парных сравнений с целью устранения метрических нарушений. Предполагается, что элементы множества, представленные матрицей парных сравнений, требуется погрузить в некоторое метрическое пространство, например, евклидово.

Как известно, в общем случае метрические нарушения на данном множестве возникают при нарушении неравенства треугольника на некоторых тройках его элементов. Более

Работа выполнена при финансовой поддержке РФФИ, проект № 13-07-00010.

жестким условием является невыполнение теоремы косинусов, когда углы в треугольнике не соответствуют длинам его сторон. В этом случае соответствующая матрица скалярных произведений тройки элементов оказывается неположительно определенной.

Следует отметить, что развиваемый нами подход имеет свой аналог в задаче экспертного оценивания, когда ранжируют элементы множества. В такой задаче часто требуется представить результат экспертизы в виде парных сравнений, причем от экспертов принято не требовать транзитивности их индивидуальных мнений. При этом исправление нарушений возлагается на метод построения ранжировки. Ранжирование означает проведение измерений в т.н. менее мощных ранговых шкалах.

Ограниченность допустимых операций в ранговых шкалах позволяет ввести простые условия транзитивности отношений, чтобы получить «сверхтранзитивную» матрицу парных сравнений, для которой выполняются условия Льюса [4, 5]. В нашем же случае измерения проводятся в более мощных, чем ранговые или интервальные, шкалах, что предполагает возможность выполнения всех обычных преобразований результатов измерений.

С другой стороны, также следует отметить, что развиваемый нами подход отличается задачи шкалирования [6, 7]. В задаче шкалирования требуется восстановить содержательно интерпретируемое пространство «стимулов». Дополнительно важным требованием обычно является минимизация его размерности. В этой связи в задаче шкалирования возникает проблема «аддитивной константы». Требуется добавить к результатам парных сравнений такую константу, чтобы матрица парных сравнений в наибольшей степени походила на матрицу скалярных произведений и имела, по возможности, наибольшее число нулевых и, если есть, то небольших по модулю отрицательных собственных значений [6]. Тогда размерность пространства стимулов определяется на основе дискретного разложения Карунена-Лоэва, которое сохраняет до 80% дисперсии исходных данных [8].

Заметим, в методе Карунена-Лоэва степень изменения значений элементов матрицы парных сравнений никак не контролируется. Кроме того, в нашем случае не требуется явного восстановления неизвестного нам признакового пространства.

Следует также отметить, что, просто восстанавливая метрическую конфигурацию элементов множества в соответствии с принципом наименьших искажений [9], мы неизбежно получаем матрицы парных сравнений с близким к нулю положительным детерминантом. Поэтому, вообще говоря, мы рискуем дополнительно получить и плохо обусловленную матрицу. Проблема плохо обусловленных матриц хорошо известна. Для обращения таких матриц разработаны весьма продвинутые вычислительные алгоритмы. В нашем подходе предполагается, что регулируемая коррекция некоторых элементов матрицы парных сравнений позволит достичь не только ее положительной определенности, но и некоторого (при более глубокой коррекции) допустимого уровня «плохой определенности».

В данной работе показано, что метрические нарушения возникают не только на тройках, но и на подмножествах, содержащих больше элементов. Такие нарушения во взаимном расположении элементов относительно друг друга также приводят к неположительной определенности матрицы парных сравнений, даже если на всех тройках элементов метрических нарушений нет. Предложены алгоритмы коррекции матриц парных сравнений. Проблема допустимого уровня плохой определенности здесь не рассматривается.

Коррекция метрических нарушений в треугольнике методом восстановления по элементам строки

Пусть множество объектов представлено парными сравнениями в виде симметричной матрицы $S(n,n)$ с элементами $s_{ij}, i = 1,\dots,n,\ \ j = 1,\dots,n$, где n – число объектов.

Недиагональные элементы принимают значения $-1 < s_{ij} < 1$ или $0 \leqslant s_{ij} < 1$. Если данная матрица положительно определена, то ее элементы рассматриваются как нормированные скалярные произведения между соответствующими векторами в метрическом пространстве и представлены косинусами углов между ними. Тогда в тройке векторов с индексами i, j и k сходство первого с самим собой $s_{ii} = 1$, с двумя остальными $s_{ij} = \cos\alpha$ и $s_{ik} = \cos\beta$.

Значение s_{ij} определяет все положения вектора с индексом j относительно вектора с индексом i как гиперконус, опирающийся на гиперокружность с центром на оси вектора с индексом i. Аналогичный смысл имеет и значение s_{ik}. Тогда относительно фиксированного вектора с индексом i возможные значения s_{jk} определены всеми парами векторов с индексами j и k, концы которых расположены на соответствующих гиперокружностях. Эти векторы наиболее близки друг к другу $c_1^{(ijk)} = s_{jk} = \cos(\beta-\alpha)$, когда расположены на одной линии по одну сторону от вектора с индексом i , и наименее близки $c_2^{(ijk)} = s_{jk} = \cos(\beta+\alpha)$, когда расположены от него по разные стороны. Если величина сходства понимается как значение неотрицательной функции близости, то $\beta + \alpha \leqslant \pi/2$. По формулам преобразования косинуса разности и суммы аргументов получим $c_{1,2}^{(ijk)} = s_{ij}s_{ik} \pm \sqrt{(1 - s_{ij}^2)(1 - s_{ik}^2)}$. Метрическое нарушение означает нарушение диапазона $c_2^{(ijk)} \leqslant s_{jk} \leqslant c_1^{(ijk)}$.

Коррекция методом восстановления по элементам строки заключается в следующем. Рассмотрим в матрице близостей S строку i, которая определяет значения сходства $s_{ij}, j = 1, \ldots, n$ вектора с индексом i с другими векторами. Относительно данного вектора все остальные образуют соответствующие гиперконусы, опирающиеся на соответствующие гиперокружности. Эти гиперокружности определяют все возможные диапазоны значений парных близостей остальных векторов множества между собой относительно данного вектора с индексом i. Поэтому по элементам каждой строки i матрицы близостей можно восстановить матрицы близостей $S^{(i)}, i = 1, \ldots, n$ с элементами в диапазонах $c_2^{(ijk)} \leqslant s_{jk}^{(i)} \leqslant c_1^{(ijk)}$.

Матрицы $S^{(i)}$ являются симметричными, с единичными главными диагоналями. В каждой из них строка i совпадает с такой же строкой в исходной матрице S . Если в матрице $S^{(i)}$ все элементы определены в соответствии с их диапазонами, то любая тройка ее элементов удовлетворяет теореме косинусов.

В общем случае для каждого значения s_{jk} можно определить не более n диапазонов для значений $s_{jk}^{(i)}$, попадание в которые не нарушит метрические соотношения. Сформируем матрицу близостей S^*, в которой для каждого ее элемента s_{jk}^* определен единственный диапазон $c_2^{*(jk)} \leqslant s_{jk}^* \leqslant c_1^{*(jk)}$, где $c_1^{*(jk)} = \min_i c_1^{(ijk)}$ и $c_2^{*(jk)} = \max_i c_2^{(ijk)}$. Для минимальной коррекции нужно выбрать, например, то значение $c_1^{*(ij)}$ или $c_2^{*(ij)}$, к которому исходное значение s_{ij} ближе всего. В итоге получим скорректированную матрицу $\tilde{S}$, в которой нет метрических нарушений на тройках элементов исходной матрицы близостей S.

Тем не менее скорректированная матрица $\tilde{S}$ снова может оказаться неположительно определенной. Это означает, что метрические нарушения также возникли на подмножествах, содержащих более трех элементов.

Коррекция нормированной матрицы близостей

Ранее нами была предложена процедура регулируемой коррекции нормированной матрицы парных близостей для устранения ее отрицательной определенности [9]. Для нормированной матрицы было показано, что метрические нарушения возникают при невыполнении теоремы косинусов, как на тройках, так и на подмножествах, содержащих большее число элементов.

В последнем случае элементы множества поочередно погружаются в специальным образом построенное координатное пространство, где корректное положение очередного добавляемого элемента определяется относительно всех ранее добавленных элементов радиусом соответствующей гиперсферы всех возможных его положений. Метрическое нарушение означает невозможность построения такой гиперсферы из-за того, что ее радиус оказывается комплексным числом.

В такой процедуре главный минор нормированной матрицы парных близостей уменьшается, начиная с единицы, оставаясь положительным, при добавлении очередного элемента множества. Если на множестве элементов возникают метрические нарушения, то главный минор текущего размера оказывается знакопеременным, постепенно уменьшаясь по модулю. Отрицательность очередного главного минора означает, что очередной добавленный элемент множества внес метрическое нарушение. Его следует исправить, корректируя значения парных сравнений нового элемента с предыдущими элементами до получения неотрицательного главного минора текущего размера.

Коррекция матрицы близостей может быть выполнена двумя способами. В первом случае корректируются все парные сравнения объекта, внесшего метрическое нарушение (вектор парных сравнений), с другими объектами. Во втором случае в матрице близостей корректируется подходящий одиночный элемент, соответствующий только одному сравнению объекта, внесшего метрическое нарушение, с некоторым другим объектом.

Коррекция вектора парных сравнений. Элементы множества просматриваются в некотором порядке, формируя множество уже просмотренных элементов, представленных на k–ом шаге главным минором $S_k = S(k,k)$, $k = 1, \ldots, n$ матрицы парных сравнений $S(n,n)$. Если минор S_k отрицателен, то считается, что именно k–й объект внес метрическое нарушение. Этот объект представлен своими парными сравнениями с предыдущими $k-1$ объектами и формирует k–ю строку и k–й столбец главного минора S_k.

Замена вектора сравнений этого элемента с предыдущими на нулевой вектор-орт единичной длины не изменит предыдущего значения детерминанта $S_k^{ort} = S_{k-1} > 0$. Действительно, в миноре S_k^{ort} такой элемент представлен нулевыми строкой и столбцом с единицей на главной диагонали. Вычисление этого минора на основе разложения по элементам последней строки заключается лишь в вычислении положительного минора S_{k-1}. Но орт-вектор парных сравнений слишком далек от вектора, вызвавшего нарушение.

Поэтому орт-вектор следует развернуть в направлении исходного вектора до положения, когда новый вектор сравнений будет не слишком сильно отличаться от исходного, а минор S_k окажется положительным.

Процедура корректировки заключается в следующем. Для заданного порога $\varepsilon > 0$ отклонения от нуля детерминант S'_k варьируется в пределах от $P_1 = S_k < 0$ до $P_2 = S_k^{ort} > 0$, где $S'_k = (P_1 + P_2)/2$. Если $S'_k \leqslant 0$, то $P_1 = S'_k$. Иначе, если $S'_k > 0$, то $P_2 = S'_k$. Если $0 \leqslant S'_k \leqslant \varepsilon$, то остановить процедуру.

Коррекция одиночных элементов. Очевидно, что одиночные коррекции уменьшают искажения, вносимые в результаты исходных парных сравнений. Далее, там, где это не вызывает затруднений, для представления значения некоторого минора будем применять такое же обозначение, как и для самого минора.

Раскроем главный минор S_k по элементам k−й строки

$$S_k = \sum_{p=1}^{k} (-1)^{k+p} s_{kp} \left(M_k\right)_p^k = \sum_{p=1}^{k-1} (-1)^{k+p} s_{kp} \left(M_k\right)_p^k + (-1)^{k+k} s_{kk} \left(M_k\right)_k^k,$$

где $\left(M_k\right)_p^k$ является дополнительным минором, полученным из главного минора S_k при удалении из него k−ой строки и p−го столбца, а второе слагаемое является предыдущим главным минором: $(-1)^{k+k} s_{kk} \left(M_k\right)_k^k = \left(M_k\right)_k^k = S_{k-1}$.

Далее раскроем миноры $\left(M_k\right)_p^k$ по элементам k−го столбца, сохранив индексацию строк и столбцов относительно исходного минора S_k. Получим

$$S_k = S_{k-1} + \sum_{p=1}^{k-1} (-1)^{k+p} s_{kp} \Big(\sum_{q=1}^{k-1} (-1)^{(q+k)-1} s_{qk} \big((M_k)_p^k \big)_k^q \Big).$$

Предполагая, что корректируются симметричные элементы $s_{kl} = s_{lk}$, представим главный минор S_k как функцию от их значений $S_k(s_{kl})$, $l = 1, \ldots, k-1$. Выполнив преобразования, получим условие коррекции каждого из элементов:

$$S_k(s_{kl}) = A_l s_{kl}^2 + B_l s_{kl} + C_l > 0,$$

$$A_l = (-1)^{2k+2l-1} \big((M_k)_l^k \big)_k^l = -\big((M_k)_l^k \big)_k^l,$$

$$B_l = 2 \sum_{q=1,\, q \neq l}^{k-1} (-1)^{2k+q+l-1} s_{qk} \big((M_k)_l^k \big)_k^q,$$

$$C_l = S_{k-1} + \sum_{p=1,\, p \neq l}^{k-1} \; \sum_{q=1,\, q \neq l}^{k-1} (-1)^{2k+p+q-1} s_{kp} s_{qk} \big((M_k)_p^k \big)_k^q .$$

Из решения следует, что коррекция симметричных элементов $s_{kl} = s_{lk}$ возможна в диапазоне $c_2^{(l)} \leqslant s_{kl} \leqslant c_1^{(l)}$, где $c_{1,2}^{(l)} = \frac{1}{2A_l} \left(-B_l \pm \sqrt{B_l^2 - 4A_l C_l} \right)$ при $B_l^2 - 4A_l C_l > 0$.

Если возможна коррекция несколько элементов, то следует выбрать тот, для которого интервал $c_1^{(l)} - c_2^{(l)}$ оказался наибольшим, и взять значение $s_{kl} = \left(c_1^{(l)} + c_2^{(l)} \right)/2$.

В этом случае будет получено максимальное положительное значение главного минора S_k, а при добавлении очередного элемента главный минор уменьшится $S_{k+1} < S_k$ в наименьшей степени.

Если коррекция одиночных элементов невозможна, то придется корректировать все парные сравнения (вектор) объекта, внесшего метрическое нарушение. В общем случае возникает задача оптимальной коррекции нескольких элементов: пар, троек и т.д. Но здесь мы ее не рассматриваем.

Оптимальная последовательность корректировок. В общем случае элементы данного множества могут быть просмотрены и в другом порядке. Может оказаться и так, что другая последовательность элементов данного множества приведет к меньшей величине суммарных искажений значений элементов исходной матрицы парных сравнений.

Пусть k–й объект внес метрическое нарушение. Пусть его коррекция выполняется одним из способов, описанных выше. Будем размещать его поочередно $q = 1, \ldots, k$ на всех местах и, выполняя каждый раз корректировку, найдем минор $S'_{k(q)}$, который в наименьшей степени отличается от минора S_k. Приняв $S_k = S'_{k(q)}$, перейдем к следующему $k+1$ объекту.

Заметим, что перестановка объекта, внесшего метрическое нарушение, может вызвать дополнительные нарушения и от других объектов. Поэтому задача поиска наименьших искажений исходной матрицы является комбинаторной, требуя просмотра всех перестановок элементов множества в общем случае.

Локализация отрицательных собственных значений. Известно, что одновременная перестановка строк и столбцов матрицы $S(n,n)$ не изменяет ее собственных значений. Согласно закону инерции квадратичных форм [10] также известно, что число смен знаков при просмотре главных миноров $S_k, k = 1, \ldots, n$ совпадает с числом отрицательных собственных значений матрицы $S(n,n)$.

Определим такой порядок просмотра элементов множества, чтобы смены знаков главных миноров $S_k, k = 1, \ldots, n$ происходили преимущественно в конце последовательности. Пусть матрица $S(n,n)$ ранга n имеет v отрицательных собственных значений. Тогда в идеальном случае соответствующая перестановка элементов множества определит такой порядок просмотра его элементов, при котором главный минор $S_{n-v+1} < 0$ впервые окажется отрицательным, а знаки последующих $v-1$ миноров будут чередоваться. В этом случае нужно будет скорректировать парные сравнения не более чем v объектов. В этом смысле отрицательные собственные значения окажутся локализованными в неположительно определенной матрице парных сравнений.

Будем удалять из матрицы парных сравнений $S(n,n)$ очередные строку и столбец, вычисляя дополнительные миноры $(M_n)_i^i$, $i = 1, \ldots, n$, и возвращать их на место. Определитель S_n матрицы $S(n,n)$ равен произведению ее собственных значений. Если этот определитель отрицателен, то имеется нечетное число отрицательных собственных значений, если положителен — то четное число. Найдем строку и столбец с номером i_1, для которых дополнительный минор $(M_n)_{i_1}^{i_1}$ сменит знак по сравнению с главным минором $M_n = S_n$ и окажется максимальным по модулю их всех возможных при $i = 1, \ldots, n$. Тогда при перемещении строки и столбца с номером i_1 на последнее место с номером n окажется, что последняя смена знака произойдет на последнем миноре $M_n = S_n$.

Продолжим этот процесс, рассматривая главные миноры $S_k,\ k = n-1, \ldots, 1$ и находя соответствующие дополнительные миноры $(M_k)_{i_q}^{i_q}$, $1 \leqslant i_q \leqslant k$. Для локализации всех отрицательных собственных значений может потребоваться более чем v шагов, так как на некоторых из них смены знака минора может не происходить. В этих случаях в очередном главном миноре S_k просто ищется соответствующий максимальный дополнительный минор $(M_k)_{i_q}^{i_q}$ без смены его знака. Пусть u – число всех дополнительных шагов, когда до локализации всех v собственных значений не происходило смены знака соответствующего главного минора. Это означает, что получена матрица $S(n,n)$, у которой придется корректировать не более $v+u$ последних элементов множества в последовательности.

Таким образом, группировка в конце последовательности объектов, предположительно вносящих метрические искажения, позволяет уменьшить дополнительные нарушения, которые могут возникнуть от других объектов в процессе коррекции.

Коррекция ненормированной матрицы близостей

Пусть множество объектов представлено взаимными ненормированными близостями в виде матрицы $S'(n,n)$ с элементами $s'_{ij}, i = 1, \ldots, n,\ j = 1, \ldots, n$, где n–число объектов.

По теореме косинусов получим $s'_{ij} = (d_{0i}^2 + d_{0j}^2 - d_{ij}^2)/2$, считая элементы матрицы $S'(n,n)$ скалярными произведениями относительно начала координат как объекта с индексом 0, где d_{ij} – расстояние между объектами с индексами i и j, d_{0i} – расстояние объекта с индексом i до начала координат. Тогда $s'_{ii} = d_{0i}^2$. Таким образом, ненормированная матрица близостей отражает конфигурацию элементов множества в метрическом пространстве, представляя их расстояния до начала координат.

Если в матрице $S'(n,n)$ имеются метрические нарушения, то их можно выявить и скорректировать, как описано выше, предварительно приведя матрицу $S'(n,n)$ к нормированному виду преобразованием $s_{ij} = s'_{ij}/\sqrt{s'_{ii}s'_{jj}}$. К сожалению, в нормированной матрице $S(n,n)$ не сохраняется исходная конфигурация объектов в пространстве, так как $s_{ii} = 1$ и «нормированные» объекты расположены на гиперсфере единичного радиуса.

Считая, что расстояния до начала координат не изменились, восстановим после корректировки ненормированную матрицу близостей $S'(n,n)$ с элементами $s'_{ij} = s_{ij}d_{0i}d_{0j}$.

В общем случае некоторые расстояния до начала координат могут измениться. Данный случай здесь рассматривать не будем.

Коррекция матрицы различий

Пусть множество объектов представлено взаимными различиями в виде матрицы $D(n,n)$ с элементами d_{ij}, $i = 1,\dots,n$, $j = 1,\dots,n$, где n – число объектов.

Будем считать различия расстояниями. Чтобы получить матрицу скалярных произведений $S(n,n)$, нужно назначить начало координат, относительного которого можно вычислить скалярные произведения. Его можно назначить в любом месте координатного пространства, например, в центре «тяжести» множества по методу Торгерсона [7].

Для размещения начала координат в центре тяжести множества объектов по методу Торгерсона получим новый объект с индексом 0, представленный своими расстояниями до остальных объектов с индексами $i = 1,\dots,n$ следующим образом:

$$d_{0i}^2 = \frac{1}{n}\sum_{p=1}^{n} d_{ip}^2 - \frac{1}{2n^2}\sum_{p=1}^{n}\sum_{q=1}^{n} d_{pq}^2. \tag{1}$$

Но иногда такое размещение начала координат не совсем удобно, особенно, если в исходной матрице парных сравнений имеются метрические нарушения. В этом случае некоторые значения (1) отрицательны, а соответствующие расстояния до начала координат оказываются комплексными.

Нужно так выбрать начало координат как новый объект, чтобы он был представлен своими действительными расстояниями до всех объектов исходного множества. Такой новый объект должен лежать вне выпуклой оболочки, образованной данным множеством объектов. Тогда, считая его началом координат, можно определить относительно него все скалярные произведения.

По теореме косинусов элементы матрицы нормированных скалярных произведений $S(n,n)$ множества объектов относительно нового объекта с индексом 0 как нового центра координат вычисляются как

$$s_{ij} = \frac{1}{2d_{0i}d_{0j}}(d_{0i}^2 + d_{0j}^2 - d_{ij}^2). \tag{2}$$

Рассмотрим выражение (1). Можно увидеть, что второе слагаемое определяет дисперсию (разброс) элементов как отклонений положения объектов от начала координат в центре тяжести множества:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} d_{0i}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n}\sum_{p=1}^{n} d_{ip}^2 - \frac{1}{2n^2}\sum_{p=1}^{n}\sum_{q=1}^{n} d_{pq}^2\right) =$$
$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{p=1}^{n} d_{ip}^2 - \frac{n}{n}\left(\frac{1}{2n^2}\sum_{p=1}^{n}\sum_{q=1}^{n} d_{pq}^2\right) = \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{p=1}^{n} d_{ip}^2.$$

Рассмотрим первое слагаемое в (1). Рассмотрим расстояния $d_{ip}, p = 1, \dots, n$ от объекта с индексом i до остальных объектов с индексами p как компоненты соответствующего вектора в некотором n–мерном метрическом пространстве, которое нам удобно назвать «вторичным». Тогда величина $\sum\limits_{p=1}^{n} d_{ip}^2$ представляет собой квадрат нормы этого вектора, т.е. квадрат расстояния от начала координат, а первое слагаемое из (1) представляет собой средневзвешенный квадрат этой нормы.

Таким образом, начало координат в таком вторичном пространстве будет представлено как объект с индексом 0^0 своими расстояниями до остальных объектов $d_{0^0 i}^2, i = 1, \dots, n$, инвариантными относительно размера множества.

Тогда начало координат по методу Торгерсона будет представлено как объект с индексом 0 своими расстояниями $d_{0i}^2 = d_{0^0 i}^2 - \Delta$, где $\Delta = \sigma^2$.

Легко увидеть, что изменение константы Δ позволит получить начало координат не в центре тяжести множества.

При $\Delta = 0$ начало координат как объект с индексом 0 будет максимально удалено от центра тяжести множества. Условие $\Delta = 0$ можно понимать как предельно маленький разброс элементов множества по сравнению с расстояниями до начала координат. Очевидно, что в этом случае скалярные произведения будут близки к единице, если получившиеся расстояния до начала координат окажутся значительными. Если $\Delta < 0$, то это свойство только усилится.

При $\Delta > \sigma^2$ обязательно возникнет неметрическая конфигурация, когда не удастся получить, согласно (2), корректный вид матрицы нормированных скалярных произведений $S(n,n)$. Это произойдет, когда некоторые из расстояний $d_{0i}^2 = d_{0^0 i}^2 - \Delta$ до начала координат окажутся нулевыми или отрицательными. Условие $\Delta > \sigma^2$ можно понимать как увеличенный по сравнению с реальным разброс элементов множества. В итоге, допустимые значения величины Δ находятся в интервале $0 \leqslant \Delta \leqslant \sigma^2$.

В общем случае метричность конфигурации определяется наличием неотрицательных собственных чисел матрицы нормированных скалярных произведений относительно назначенного начала координат. Поэтому следует так разместить начало координат, чтобы матрица нормированных скалярных произведений относительно него была бы неотрицательно или положительно определена.

Если в исходной матрице расстояний имелись нарушения метрики, то полученную матрицу скалярных произведений $S(n,n)$ необходимо откорректировать.

После исправления метрических нарушений получим положительно определенную матрицу скалярных произведений $S(n,n)$.

Далее восстановим ненормированную матрицу $S'(n,n)$ скалярных произведений с элементами $s'_{ij} = s_{ij} d_{0i} d_{0j}$ и матрицу расстояний $D(n,n)$, считая, что расстояния от объектов до начала координат не изменились, где $d_{ij}^2 = s'_{ii} + s'_{jj} - 2s'_{ij} = d_{0i}^2 + d_{0j}^2 - 2s_{ij} d_{0i} d_{0j}$.

Эксперименты

Эксперименты были проведены на трех матрицах расстояний между некоторыми населенными пунктами Тульской области. Первая матрица была получена путем непосредственного измерения по карте, в результате которого некоторые расстояния были намеренно искажены:

$$D_C = \begin{array}{c@{}c} & \begin{matrix} \text{Т} & \text{А} & \text{Н} & \text{Б} & \text{Я} & \text{Ар} & \text{Е} \end{matrix} \\ \begin{matrix} \text{Т} \\ \text{А} \\ \text{Н} \\ \text{Б} \\ \text{Я} \\ \text{Ар} \\ \text{Е} \end{matrix} & \begin{pmatrix} 0 & 67 & 62 & 123 & 50 & 105 & 150 \\ 67 & 0 & 144 & 149 & 45 & 128 & 218 \\ 62 & 144 & 0 & 180 & 116 & 164 & 135 \\ 123 & 149 & 180 & 0 & 160 & 75 & 231 \\ 50 & 45 & 116 & 160 & 0 & 142 & 186 \\ 105 & 128 & 164 & 75 & 142 & 0 & 202 \\ 150 & 218 & 135 & 231 & 186 & 202 & 0 \end{pmatrix} \end{array}.$$

В данной матрице имеются нарушения неравенства треугольника на некоторых тройках элементов. Для трех городов: *Тула*, *Алексин* и *Новомосковск*, суммарное расстояние от *Тулы* до *Алексина* и *Новомосковска* (129 км) оказалось меньше расстояния от *Новомосковска* до *Алексина* (144 км). Также суммарное расстояние от *Тулы* до *Алексина* и *Ефремова* (217 км) оказалось меньше расстояния от *Алексина* до *Ефремова* (218 км). Наконец, суммарное расстояние от *Тулы* до *Новомосковска* и *Ясногорска* (112 км) меньше расстояние от *Новомосковска* до *Ясногорска* (116 км).

Вторая матрица была получена с помощью сервиса Google*:

$$D_G = \begin{array}{c@{}c} & \begin{matrix} \text{Т} & \text{А} & \text{Н} & \text{Б} & \text{Я} & \text{Ар} & \text{Е} \end{matrix} \\ \begin{matrix} \text{Т} \\ \text{А} \\ \text{Н} \\ \text{Б} \\ \text{Я} \\ \text{Ар} \\ \text{Е} \end{matrix} & \begin{pmatrix} 0 & 48 & 42 & 106 & 31 & 81 & 121 \\ 48 & 0 & 88 & 99 & 40 & 90 & 165 \\ 42 & 88 & 0 & 140 & 55 & 110 & 104 \\ 106 & 99 & 140 & 0 & 126 & 36 & 150 \\ 31 & 40 & 55 & 126 & 0 & 107 & 150 \\ 81 & 90 & 110 & 36 & 107 & 0 & 116 \\ 121 & 165 & 104 & 150 & 150 & 116 & 0 \end{pmatrix} \end{array}$$

и третья матрица была получена с помощью сервиса Yandex:

$$D_Y = \begin{array}{c@{}c} & \begin{matrix} \text{Т} & \text{А} & \text{Н} & \text{Б} & \text{Я} & \text{Ар} & \text{Е} \end{matrix} \\ \begin{matrix} \text{Т} \\ \text{А} \\ \text{Н} \\ \text{Б} \\ \text{Я} \\ \text{Ар} \\ \text{Е} \end{matrix} & \begin{pmatrix} 0 & 50.3 & 49.7 & 107 & 31.8 & 81 & 121 \\ 50.3 & 0 & 98.6 & 99.9 & 39.7 & 91.1 & 166 \\ 49.7 & 98.6 & 0 & 144 & 67.2 & 111 & 94.9 \\ 107 & 99.9 & 144 & 0 & 126 & 36.2 & 149 \\ 31.8 & 39.7 & 67.2 & 126 & 0 & 106 & 150 \\ 81 & 91.1 & 111 & 36.2 & 106 & 0 & 115 \\ 121 & 166 & 94.9 & 149 & 150 & 115 & 0 \end{pmatrix} \end{array}.$$

В этих матрицах неравенства треугольника выполняются.

*Здесь и далее Т - Тула, А - Алексин, Н - Новомосковск, Б - Белёв, Я - Ясногорск, Ар - Арсеньево, Е - Ефремов (исправления внесены 16/01/15 по запросу автора)

Для всех матриц расстояний были получены матрицы нормированных скалярных произведений и их собственные значения:

$$S_C = \begin{pmatrix} 1 & 0.8611 & 0.8933 & 0.5631 & 0.9174 & 0.6085 & 0.5237 \\ 0.8611 & 1 & 0.3659 & 0.4244 & 0.9354 & 0.5107 & -0.0073 \\ 0.8933 & 0.3659 & 1 & 0.1613 & 0.5615 & 0.2038 & 0.6472 \\ 0.5631 & 0.4244 & 0.1613 & 1 & 0.2981 & 0.8633 & -0.0060 \\ 0.9174 & 0.9354 & 0.5615 & 0.2981 & 1 & 0.3581 & 0.2505 \\ 0.6085 & 0.5107 & 0.2038 & 0.8633 & 0.3581 & 1 & 0.1614 \\ 0.5237 & -0.0073 & 0.6472 & -0.0060 & 0.2505 & 0.1614 & 1 \end{pmatrix},$$

$$\boldsymbol{\lambda}^{\mathsf{T}}_{S_C} = (4.075 \;\; 1.571 \;\; 1.047 \;\; 0.307 \;\; 0.112 \;\; 0.009 \;\; -0.121),$$

$$S_G = \begin{pmatrix} 1 & 0.8468 & 0.8826 & 0.3566 & 0.9454 & 0.4995 & 0.3573 \\ 0.8468 & 1 & 0.5153 & 0.5071 & 0.9002 & 0.4836 & -0.1427 \\ 0.8826 & 0.5153 & 1 & -0.0139 & 0.8083 & 0.2159 & 0.5784 \\ 0.3566 & 0.5071 & -0.0139 & 1 & 0.1843 & 0.9522 & 0.1797 \\ 0.9454 & 0.9002 & 0.8083 & 0.1843 & 1 & 0.2606 & 0.0577 \\ 0.4995 & 0.4836 & 0.2159 & 0.9522 & 0.2606 & 1 & 0.4556 \\ 0.3573 & -0.1427 & 0.5784 & 0.1797 & 0.0577 & 0.4556 & 1 \end{pmatrix},$$

$$\boldsymbol{\lambda}^{\mathsf{T}}_{S_G} = (3.994 \;\; 1.681 \;\; 1.326 \;\; 0.008 \;\; -0.001 \;\; -0.004 \;\; -0.005),$$

$$S_Y = \begin{pmatrix} 1 & 0.8392 & 0.8406 & 0.3568 & 0.9428 & 0.5049 & 0.3487 \\ 0.8392 & 1 & 0.4254 & 0.5115 & 0.9055 & 0.4850 & -0.1493 \\ 0.8406 & 0.4254 & 1 & -0.0349 & 0.7264 & 0.2283 & 0.6547 \\ 0.3568 & 0.5115 & -0.0349 & 1 & 0.1995 & 0.9536 & 0.1870 \\ 0.9428 & 0.9055 & 0.7264 & 0.1995 & 1 & 0.2830 & 0.0530 \\ 0.5049 & 0.4850 & 0.2283 & 0.9536 & 0.2830 & 1 & 0.4572 \\ 0.3487 & -0.1493 & 0.6547 & 0.1870 & 0.0530 & 0.4572 & 1 \end{pmatrix},$$

$$\boldsymbol{\lambda}^{\mathsf{T}}_{S_Y} = (3.94 \;\; 1.639 \;\; 1.418 \;\; 0.007 \;\; 0.002 \;\; -0.001 \;\; -0.005).$$

Все нормированные матрицы были получены при $\Delta = 0$ относительно начала координат, как объекта с индексом 0, вынесенного за пределы выпуклой оболочки данного множества элементов на соответствующие расстояния:

$$\boldsymbol{d}^{\mathsf{T}}_{0_C} = (92.19 \;\; 127.26 \;\; 128.48 \;\; 148.48 \;\; 118.53 \;\; 131.42 \;\; 176.07),$$

$$\boldsymbol{d}^{\mathsf{T}}_{0_G} = (73.16 \;\; 90.060 \;\; 88.680 \;\; 107.11 \;\; 88.970 \;\; 86.980 \;\; 126.003),$$

$$\boldsymbol{d}^{\mathsf{T}}_{0_Y} = (74.32 \;\; 92.320 \;\; 91.630 \;\; 107.94 \;\; 90.010 \;\; 86.980 \;\; 124.860).$$

На всех тройках объектов, представленных парными близостями в матрице S_C, для которых в матрице D_C было нарушено неравенство треугольника:

$$S_{--} = \begin{pmatrix} 1 & 0.8611 & 0.8933 \\ 0.8611 & 1 & 0.3659 \\ 0.8933 & 0.3659 & 1 \end{pmatrix}, \; \boldsymbol{\lambda} = \begin{pmatrix} 2.437 \\ 0.634 \\ -0.071 \end{pmatrix},$$

$$S_{--} = \begin{pmatrix} 1 & 0.8611 & 0.5237 \\ 0.8611 & 1 & -0.0073 \\ 0.5237 & -0.0073 & 1 \end{pmatrix}, \; \boldsymbol{\lambda} = \begin{pmatrix} 2.005 \\ 1.007 \\ -0.011 \end{pmatrix},$$

$$S_{--} = \begin{pmatrix} 1 & 0.8933 & 0.9174 \\ 0.8933 & 1 & 0.5615 \\ 0.9174 & 0.5615 & 1 \end{pmatrix}, \ \boldsymbol{\lambda} = \begin{pmatrix} 2.592 \\ 0.439 \\ -0.030 \end{pmatrix}$$

были скорректированы близости по элементам строк этих подматриц.

Выполним коррекцию близостей по элементам первой строки матрицы S_{--}. В матрице $S^{(1)}$ элементы первой строки не изменятся, элементы второй строки $s_{21}^{(1)}$ и $s_{22}^{(1)}$ также не изменятся из-за симметрии, а для элемента $s_{23}^{(1)}$ определен диапазон $0.5407 \leqslant s_{23}^{(1)} \leqslant 0.9977$:

$$c_{1,2}^{(123)} = s_{12}s_{13} \pm \sqrt{(1-s_{12}^2)(1-s_{13}^2)} = 0.7692 \pm 0.2285\ .$$

Значение $s_{23} = 0.3659$ за пределами диапазона вносит метрическое искажение относительно элементов первой строки исходной матрицы. Элементы третьей строки уже определены в силу симметрии. Можно убедиться, что для элементов $s_{31}^{(1)}$ и $s_{33}^{(1)}$ их диапазоны содержат исходные значения s_{31} и $s_{33} = 1$:

$$c_{1,2}^{(131)} = s_{13}s_{11} \pm \sqrt{(1-s_{13}^2)(1-s_{11}^2)} = s_{13} \pm 0 = s_{13},$$
$$c_{1,2}^{(133)} = s_{13}s_{13} \pm \sqrt{(1-s_{13}^2)(1-s_{13}^2)} = s_{13}^2 \pm (1-s_{13}^2),$$

а $s_{32} = 0.3659$ также находится за пределами диапазона $0.4443 \leqslant s_{32}^{(1)} \leqslant 0.9939$:

$$c_{1,2}^{(132)} = s_{13}s_{12} \pm \sqrt{(1-s_{13}^2)(1-s_{12}^2)} = 0.7692 \pm 0.2285\ .$$

В итоге, восстановление по элементам каждой из строк дает матрицы

$$S^{(1)} = \begin{pmatrix} 1 & 0.8611 & 0.8933 \\ 0.8611 & 1 & 0.5407 \div 0,9977 \\ 0.8933 & 0.5407 \div 0.9977 & 1 \end{pmatrix},$$

$$S^{(2)} = \begin{pmatrix} 1 & 0.8611 & 0.1581 \div 0.7883 \\ 0.8611 & 1 & 0.3659 \\ 0.1581 \div 0.7883 & 0.3659 & 1 \end{pmatrix},$$

$$S^{(3)} = \begin{pmatrix} 1 & -0.0914 \div 0,7452 & 0.8933 \\ -0.0914 \div 0.7452 & 1 & 0.3659 \\ 0.8933 & 0.3659 & 1 \end{pmatrix}.$$

Объединение диапазонов дает матрицу без метрических нарушений

$$\tilde{S}_{--} = \begin{pmatrix} 1 & 0.7452 & 0.7883 \\ 0.7452 & 1 & 0.5407 \\ 0.7883 & 0.5407 & 1 \end{pmatrix}, \ \boldsymbol{\lambda} = \begin{pmatrix} 2.388 \\ 0.461 \\ 0.151 \end{pmatrix}.$$

Так же были получены и две другие матрицы, где для повторяющихся пар городов *Тула-Алексин* и *Тула-Новомосковск* были взяты одни и те же значения близостей, так как они попали в соответствующие диапазоны:

$$\tilde{S}_{--} = \begin{pmatrix} 1 & 0.7452 & 0.5021 \\ 0.7452 & 1 & 0.0178 \\ 0.5021 & 0.0178 & 1 \end{pmatrix}, \ \boldsymbol{\lambda} = \begin{pmatrix} 1.907 \\ 0.984 \\ 0.110 \end{pmatrix},$$

$$\tilde{S}_{--} = \begin{pmatrix} 1 & 0.7883 & 0.8735 \\ 0.7883 & 1 & 0.6406 \\ 0.8735 & 0.6406 & 1 \end{pmatrix}, \ \boldsymbol{\lambda} = \begin{pmatrix} 2.539 \\ 0.367 \\ 0.094 \end{pmatrix}.$$

При объединении всех корректировок была получена матрица близостей, в которой нет ни одного нарушения неравенства треугольника. Тем не менее у нее снова оказалось одно отрицательное собственное значение

$$\boldsymbol{\lambda}^{\mathsf{T}}_{\tilde{S}_C} = (4.048 \ \ 1.544 \ \ 1.012 \ \ 0.237 \ \ 0.139 \ \ 0.049 \ \ -0.029).$$

Поэтому дополнительно была выполнена коррекция ее симметричных элементов s_{25} и s_{52} на основе процедуры локализации отрицательных собственных значений. Все измененные значения показаны жирным шрифтом. Полученная матрица

$$\tilde{S}_C = \begin{pmatrix} 1 & \mathbf{0.7452} & \mathbf{0.7883} & 0.5631 & \mathbf{0.8735} & 0.6085 & \mathbf{0.5021} \\ \mathbf{0.7452} & 1 & \mathbf{0.5407} & 0.4244 & \mathbf{0.8729} & 0.5107 & \mathbf{0.0178} \\ \mathbf{0.7883} & \mathbf{0.5407} & 1 & 0.1613 & \mathbf{0.6406} & 0.2038 & 0.6472 \\ 0.5631 & 0.4244 & 0.1613 & 1 & 0.2981 & 0.8633 & -0.0060 \\ \mathbf{0.8735} & \mathbf{0.8729} & \mathbf{0.6406} & 0.2981 & 1 & 0.3581 & 0.2505 \\ 0.6085 & 0.5107 & 0.2038 & 0.8633 & 0.3581 & 1 & 0.1614 \\ \mathbf{0.5021} & \mathbf{0.0178} & 0.6472 & -0.0060 & 0.2505 & 0.1614 & 1 \end{pmatrix}$$

имеет суммарное отличие от исходной 0.6282 и имеет только положительные собственные значения

$$\boldsymbol{\lambda}^{\mathsf{T}}_{\tilde{S}_C} = (4.025 \ \ 1.545 \ \ 0.986 \ \ 0.232 \ \ 0.162 \ \ 0.05 \ \ 0.000005).$$

Восстановленная матрица расстояний имеет вид

$$\tilde{D}_C = \left(\begin{array}{c|ccccccc} & \textsf{Т} & \textsf{А} & \textsf{Н} & \textsf{Б} & \textsf{Я} & \textsf{Ар} & \textsf{Е} \\ \hline \textsf{Т} & 0 & 84.9 & 79.6 & 123 & 58.8 & 105 & 152.3 \\ \textsf{А} & 84.9 & 0 & 122.6 & 149 & 62.5 & 128 & 215.4 \\ \textsf{Н} & 79.6 & 122.6 & 0 & 180 & 105.1 & 164 & 135 \\ \textsf{Б} & 123 & 149 & 180 & 0 & 160 & 75 & 231 \\ \textsf{Я} & 58.8 & 62.5 & 105.1 & 160 & 0 & 142 & 186 \\ \textsf{Ар} & 105 & 128 & 164 & 75 & 142 & 0 & 202 \\ \textsf{Е} & 152.3 & 215.4 & 135 & 231 & 186 & 202 & 0 \end{array}\right).$$

Для матрицы S_C также была выполнена корректировка без предварительного устранения нарушений неравенства треугольника. Коррекция одиночных элементов оказалась невозможна, поэтому была выполнена коррекция вектором на основе процедуры локализации отрицательных собственных значений. Полученная матрица

$$S_C = \begin{pmatrix} 1 & \mathbf{0.7434} & \mathbf{0.7712} & \mathbf{0.4861} & \mathbf{0.7919} & \mathbf{0.5253} & \mathbf{0.4521} \\ \mathbf{0.7434} & 1 & 0.3659 & 0.4244 & 0.9354 & 0.5107 & -0.0073 \\ \mathbf{0.7712} & 0.3659 & 1 & 0.1613 & 0.5615 & 0.2038 & 0.6472 \\ \mathbf{0.4861} & 0.4244 & 0.1613 & 1 & 0.2981 & 0.8633 & -0.0060 \\ \mathbf{0.7919} & 0.9354 & 0.5615 & 0.2981 & 1 & 0.3581 & 0.2505 \\ \mathbf{0.5253} & 0.5107 & 0.2038 & 0.8633 & 0.3581 & 1 & 0.1614 \\ \mathbf{0.4521} & -0.0073 & 0.6472 & -0.0060 & 0.2505 & 0.1614 & 1 \end{pmatrix}$$

имеет суммарное отличие от исходной 0.5971 и собственные значения

$$\boldsymbol{\lambda}^{\mathsf{T}}_{S_C} = (3.861 \ \ 1.569 \ \ 1.047 \ \ 0.314 \ \ 0.112 \ \ 0.096 \ \ 0.0003).$$

Восстановленная матрица расстояний имеет вид

$$D_C = \begin{array}{rl} & \begin{matrix} \text{Т} & \text{А} & \text{Н} & \text{Б} & \text{Я} & \text{Ар} & \text{Е} \end{matrix} \\ \begin{matrix} \text{Т} \\ \text{А} \\ \text{Н} \\ \text{Б} \\ \text{Я} \\ \text{Ар} \\ \text{Е} \end{matrix} & \begin{pmatrix} 0 & 85.2 & 82.1 & 131.3 & 72.4 & 114.2 & 157.6 \\ 85.2 & 0 & 144 & 149 & 45 & 129 & 218 \\ 82.1 & 144 & 0 & 180 & 116 & 164 & 135 \\ 131.3 & 149 & 180 & 0 & 160 & 75 & 231 \\ 72.4 & 45 & 116 & 160 & 0 & 142 & 186 \\ 114.2 & 128 & 164 & 75 & 142 & 0 & 202 \\ 157.6 & 218 & 135 & 231 & 186 & 202 & 0 \end{pmatrix} \end{array}.$$

Корректировка матриц S_G и S_Y также была выполнена основе процедуры локализации отрицательных собственных значений, их суммарные отличия от исходных матриц составляют 0.0605 и 0.0298. При этом корректировались как одиночные элементы, так и векторы:

$$S_G = \begin{pmatrix} 1 & \mathbf{0.8402} & \mathbf{0.8757} & \mathbf{0.3538} & \mathbf{0.9380} & \mathbf{0.4956} & \mathbf{0.3546} \\ \mathbf{0.8402} & 1 & 0.5153 & 0.5071 & 0.9002 & \mathbf{0.4798} & -0.1427 \\ \mathbf{0.8757} & 0.5153 & 1 & -0.0139 & \mathbf{0.7967} & \mathbf{0.2142} & 0.5784 \\ \mathbf{0.3538} & 0.5071 & -0.0139 & 1 & 0.1843 & \mathbf{0.9448} & 0.1797 \\ \mathbf{0.9380} & 0.9002 & \mathbf{0.7967} & 0.1843 & 1 & \mathbf{0.2585} & 0.0577 \\ \mathbf{0.4956} & \mathbf{0.4798} & \mathbf{0.2142} & \mathbf{0.9448} & \mathbf{0.2585} & 1 & \mathbf{0.4520} \\ \mathbf{0.3546} & -0.1427 & 0.5784 & 0.1797 & 0.0577 & \mathbf{0.4520} & 1 \end{pmatrix},$$

$$\boldsymbol{\lambda}_{S_G}^{\mathsf{T}} = (3.975 \;\; 1.674 \;\; 1.327 \;\; 0.014 \;\; 0.006 \;\; 0.004 \;\; 0.0003),$$

$$S_Y = \begin{pmatrix} 1 & 0.8392 & \mathbf{0.8341} & 0.3568 & 0.9428 & 0.5049 & 0.3487 \\ 0.8392 & 1 & \mathbf{0.4221} & \mathbf{0.5044} & 0.9055 & 0.4850 & -0.1493 \\ \mathbf{0.8341} & \mathbf{0.4221} & 1 & \mathbf{-0.0346} & \mathbf{0.7207} & \mathbf{0.2266} & \mathbf{0.6496} \\ 0.3568 & \mathbf{0.5044} & \mathbf{-0.0346} & 1 & 0.1995 & 0.9536 & 0.1870 \\ 0.9428 & 0.9055 & \mathbf{0.7207} & 0.1995 & 1 & 0.2830 & 0.0530 \\ 0.5049 & 0.4850 & \mathbf{0.2266} & 0.9536 & 0.2830 & 1 & 0.4572 \\ 0.3487 & -0.1493 & \mathbf{0.6496} & 0.1870 & 0.0530 & 0.4572 & 1 \end{pmatrix},$$

$$\boldsymbol{\lambda}_{S_Y}^{\mathsf{T}} = (3.931 \;\; 1.638 \;\; 1.417 \;\; 0.009 \;\; 0.004 \;\; 0.001 \;\; 0.0001).$$

Восстановленные матрицы расстояний имеют вид:

$$D_G = \begin{array}{rl} & \begin{matrix} \text{Т} & \text{А} & \text{Н} & \text{Б} & \text{Я} & \text{Ар} & \text{Е} \end{matrix} \\ \begin{matrix} \text{Т} \\ \text{А} \\ \text{Н} \\ \text{Б} \\ \text{Я} \\ \text{Ар} \\ \text{Е} \end{matrix} & \begin{pmatrix} 0 & 48.9 & 43.1 & 106.2 & 32.5 & 81.3 & 121.2 \\ 48.9 & 0 & 88 & 99 & 40 & 90.3 & 165 \\ 43.1 & 88 & 0 & 140 & 56.6 & 110.1 & 104 \\ 106.2 & 99 & 140 & 0 & 126 & 37.9 & 150 \\ 32.5 & 40 & 56.6 & 126 & 0 & 107.1 & 150 \\ 81.3 & 90.3 & 110.1 & 37.9 & 107.1 & 0 & 116.3 \\ 121.2 & 165 & 104 & 150 & 150 & 116.3 & 0 \end{pmatrix} \end{array},$$

$$D_Y = \begin{array}{c|ccccccc} & Т & А & Н & Б & Я & Ар & Е \\ \hline Т & 0 & 50.3 & 50.6 & 107 & 31.8 & 81 & 121 \\ А & 50.3 & 0 & 98.9 & 100.6 & 39.7 & 91.1 & 166 \\ Н & 50.6 & 98.9 & 0 & 144 & 67.9 & 111.1 & 95.5 \\ Б & 107 & 100.6 & 144 & 0 & 126 & 36.2 & 149 \\ Я & 31.8 & 39.7 & 67.9 & 126 & 0 & 106 & 150 \\ Ар & 81 & 91.1 & 111.1 & 36.2 & 106 & 0 & 115 \\ Е & 121 & 166 & 95.5 & 149 & 150 & 115 & 0 \end{array}.$$

Результаты работы процедуры локализации отрицательных собственных значений представлены в таблице 1.

Таблица 1: Локализация отрицательных собственных значений

Матрица	n	v	Перестановка объектов	Позиция $n-v+1$	Коррекция объекта
$\tilde{S}_C$	7	1	7 4 2 3 6 1 5	7	5
S_C	7	1	7 4 5 3 6 2 1	7	1
S_G	7	3	7 2 4 3 5 1 6	5	5 1 6
S_Y	7	2	7 1 6 2 5 4 3	6	4 3

Все рассмотренные выше матрицы скалярных произведений имеют ранг $n = 7$ и свое число v отрицательных собственных значений. Процедура локализации выявляет оптимальную перестановку объектов и позицию $n - v + 1$ первого отрицательного главного минора. В соответствии с этими позициями оказалось, что в каждой матрице корректировалось минимальное число объектов v, в точности совпадающее с соответствующим числом отрицательных собственных значений.

Заключение

Что есть истина?
(И. 18: 33, 36-38)

В данной работе рассмотрен один из возможных подходов к погружению элементов множества в метрическое пространство и проведены эксперименты по коррекции расстояний на примере некоторых населенных пунктов Тульской области.

Условием корректного погружения является отсутствие метрических нарушений во взаимном расположении элементов множества друг относительно друга.

Следует отметить, что в задаче коррекции нет речи о восстановлении каких-то «истинных» расстояний, так как условием корректировки является только восстановление положительной определенности соответствующей матрицы парных сравнений. Это видно из результата восстановления первой матрицы расстояний, в которую были внесены большие искажения. В восстановленной и теперь корректной матрице эти искажения «истинных» расстояний остались.

Показано, что суть нарушений как на тройках, так и на большем числе элементов множества, заключается, в итоге, в нарушении теоремы косинусов. Это удалось показать в явном виде [9] только для нормированной матрицы парных сравнений, имеющих смысл скалярных произведений.

В этих условиях потребовалась разработка специальных процедур для корректировки произвольных матриц парных сравнений, которые могут быть матрицами расстояний и ненормированными матрицами близостей или скалярных произведений.

Литература

[1] *Young G., Householder A. S.* Discussion of a set of points in terms of their mutual distances // *Psychometrica.* 1938. Vol. 3, no. 1. P. 19–22.

[2] *Двоенко С. Д.* Распознавание элементов множества, представленных взаимными расстояниями и близостями // *ММРО-14*. М.: МАКС Пресс, 2009. С. 112–115.

[3] *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.

[4] *Киселев Ю. В.* Оценка важности программ методом парных сравнений // *Известия АН СССР. Техническая кибернетика.* 1971. № 3. С. 41–47.

[5] *Миркин Б. Г.* Проблема группового выбора. М.: Наука, 1974. 256 с.

[6] *Терехина А. Ю.* Анализ данных методами многомерного шкалирования. М.: Наука, 1986. 168 с.

[7] *Дэйвисон М.* Многомерное шкалирование. М.: Финансы и статистика, 1988. 254 с.

[8] *Ту Дж., Гонсалес Р.* Принципы распознавания образов. М.: Мир, 1978. 411 с.

[9] *Двоенко С. Д., Пшеничный Д. О.* Об устранении отрицательных собственных значений матриц парных сравнений // *ИОИ-2012.* М.: ТОРУС ПРЕСС, 2012. С. 13–16.

[10] *Гантмахер Ф. Р.* Теория матриц. М.: Наука, 1988. 552 с.