

О КЛАССИФИКАЦИЯХ ГЕОИНФОРМАЦИОННЫХ ОБЪЕКТОВ

Миронова Ю.Н. ©

Кандидат физико-математических наук,
доцент кафедры математики и прикладной информатики,
Елабужский институт Казанского федерального университета

Аннотация

В статье рассматривается классификация геоинформационных объектов. При классификациях, относящихся к признакам числовой природы, удобно использовать матрицы признаков. Гораздо сложнее это выполнить для объектов нечисловой природы, таких, как лингвистическая переменная. Можно предложить изобразить множество значений матрицы графически, вложив их в лексикографически упорядоченный квадрат. Данный метод удобен тем, что мы в наше пространство можем поместить как числовые, так и нечисловые объекты.

Ключевые слова: классификация, лингвистическая переменная, лексикографически упорядоченный квадрат, геоинформатика.

Keywords: classification, linguistic variable, the lexicographically ordered square, Geoinformatics.

Классификация изучаемых явлений и процессов в присуща всем наукам, в том числе и геоинформатике, в которой классификация часто выступает не только методом, но и целью научного исследования. В литературных источниках наряду с термином «классификация» в близких смыслах используются термины «группировка», «распознавание образов», «диагностика», «дискриминация», «сортировка» и др. [10]

Результатом представления геоинформационного объекта в признаковом пространстве является некоторая матрица ([1], [5], [8]). Если мы имеем N объектов, обладающих каждый M признаками, то

$$X = \begin{pmatrix} o_1 \\ \dots \\ o_N \end{pmatrix} = (o^{(1)}, \dots, o^{(M)}) = \begin{pmatrix} o_1^{(1)} & \dots & o_1^{(j)} & \dots & o_1^{(M)} \\ o_2^{(1)} & \dots & o_2^{(j)} & \dots & o_2^{(M)} \\ \dots & \dots & \dots & \dots & \dots \\ o_{N-1}^{(1)} & \dots & o_{N-1}^{(j)} & \dots & o_{N-1}^{(M)} \\ o_N^{(1)} & \dots & o_N^{(j)} & \dots & o_N^{(M)} \end{pmatrix}$$

Здесь $o_i^{(j)}$ - значение j -го признака для i -го элемента нашей системы геоинформационных объектов.

Можно сопоставить две таких матрицы, и получить матрицу различий наших систем:

$$A = \begin{pmatrix} a_{11} & \dots & a_{ij} & \dots & a_{1M} \\ a_{21} & \dots & a_{2j} & \dots & a_{2M} \\ \dots & \dots & \dots & \dots & \dots \\ a_{i1} & \dots & a_{ij} & \dots & a_{iM} \\ \dots & \dots & \dots & \dots & \dots \\ a_{N1} & \dots & a_{Nj} & \dots & a_{NM} \end{pmatrix}$$

Здесь a_{ij} - результат сопоставления i -й и j -й системы геоинформационных объектов. Обычно a_{ij} означает меру различия (или сходства) объектов.

Переход к матрице A осуществляется с помощью задания метрики d (или расстояния между геоинформационными объектами [4]).

При классификациях, относящихся к признакам числовой природы, удобно использовать матрицы признаков:

$$B = \begin{pmatrix} b_{11} & \dots & b_{ij} & \dots & b_{1M} \\ b_{21} & \dots & b_{2j} & \dots & b_{2M} \\ \dots & \dots & \dots & \dots & \dots \\ b_{i1} & \dots & b_{ij} & \dots & b_{iM} \\ \dots & \dots & \dots & \dots & \dots \\ b_{N1} & \dots & b_{Nj} & \dots & b_{NM} \end{pmatrix}$$

Можно ввести расстояние между признаками, и на этом основании производить классификацию объектов.

Гораздо сложнее это выполнить для объектов нечисловой природы, таких, как лингвистическая переменная. Лингвистическая переменная отличается от числовой переменной тем, что ее значениями являются не числа, а слова или предложения в естественном или формальном языке [7].

Например, лингвистической переменной является понятие «Лес», ее значения: «лиственный», «смешанный», «хвойный». Количественные и качественные характеристики объекта можно также рассматривать как лингвистическую переменную: «очень мало», «мало», «средне», «много», «очень много», или «очень плохо», «плохо», «средне», «хорошо», «очень хорошо». То есть значения лингвистических переменных также можно записать как матрицу $m \times n$, но будут известны не все элементы матрицы, а известные значения не будут числами. Ясно, что с такой матрицей будет трудно работать. Поэтому можно предложить изобразить множество значений матрицы графически.

Рассмотрим лексикографически упорядоченный квадрат Q , описанный в работе [6].

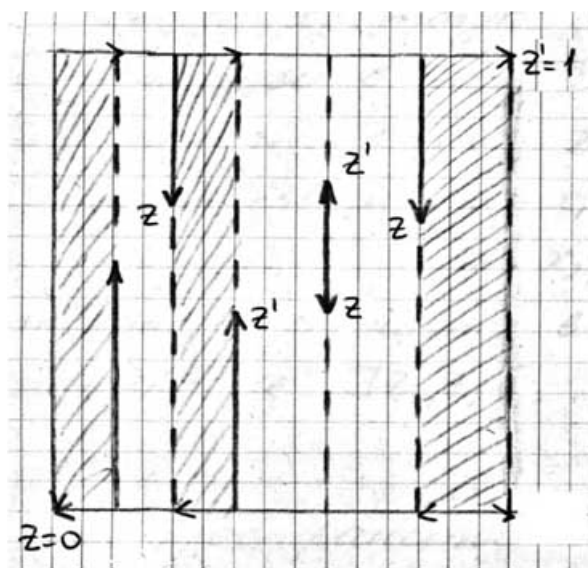


Рис. 1. Лексикографически упорядоченный квадрат [7]

Напомним определение нашего пространства.

Рассмотрим на плоскости OXY замкнутый квадрат со сторонами, параллельными осям координат и вершинами $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ и упорядочим множество всех точек $z = (x, y)$, $0 \leq x \leq 1$, $0 \leq y \leq 1$ этого квадрата в лексикографическом порядке, то есть:

$(x, y) < (x', y')$, если $x < x'$ или
если $x = x'$ и $y < y'$.

Полученные в результате такого упорядочения порядковые интервалы и полуинтервалы $[0, \alpha[$ и $] \beta, 1]$ образуют базу нашего пространства Q .

Эти интервалы имеют следующий вид: пусть даны $z_1 < z_2$, $z_1 = (x_1, y_1)$, $z_2 = (x_2, y_2)$, причем $x_1 < x_2$, тогда для любой точки z , лежащей в полосе $0 \leq y \leq 1$, $x_1 < x < x_2$, мы получим, что $z_1 < z < z_2$.

Полуинтервалы $x = x_1$, $y_1 < y \leq 1$ и $x = x_2$, $0 \leq y \leq y_2$ также содержатся в порядковом интервале $]z_1, z_2[$, если $y_1 \neq 1$, $y_2 \neq 0$.

Расположим наши лингвистические переменные по горизонтали в соответствии с их весами. По вертикали расположим их значения в соответствии с увеличением (это возможно, так как наше пространство является линейно упорядоченным).

Слева направо и снизу вверх соответствующие величины увеличиваются. Далее рассматриваем свойства получившегося объекта. (Рис. 2)

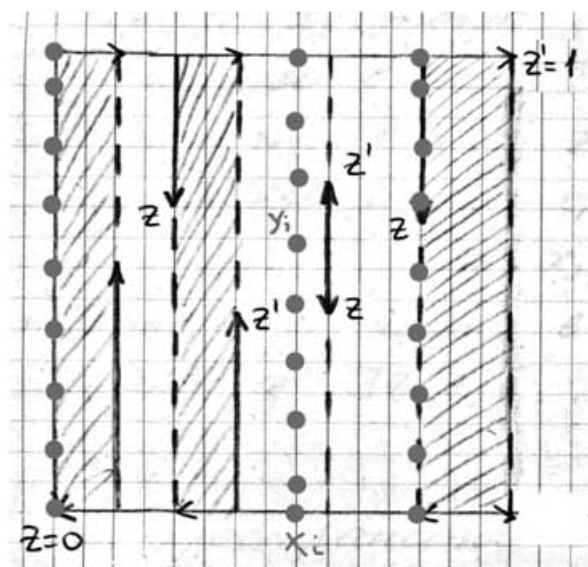


Рис. 2. Лексикографически упорядоченный квадрат со вложенной в него системой точек [7]

Данный метод удобен тем, что мы в наше пространство можем поместить как числовые, так и нечисловые объекты. Если числа b_{ij} в матрице B превышают единицу, то их можно нормировать, и уместить в сегменте $[0,1]$ как по горизонтали, так и по вертикали. Так как Q - топологическое пространство с лексикографическим порядком, то можно растягивать или сжимать соответствующие отрезки без изменения результата. Так как мы имеем только отношения «больше - меньше», то можем отказаться от метрики, и уменьшить субъективность рассмотрения объекта, то есть выявить связи, не зависящие от метрики.

Таким образом, мы можем применять некоторые математические методы ([4], [5]) в науках, где исходные данные не всегда могут быть точно выражены, например, в геоинформатике [9], статистике [10] и других науках.

Литература

1. Капралов Е.Г. Геоинформатика: в 2 кн. Кн. 1: учебник для студ. высш. учеб. заведений. / [Е.Г. Капралов, А.В. Кошкарёв, В.С. Тикунов и др.]; под ред. В.С. Тикунова. – 3-е изд., перераб. и доп. – М.: Издательский центр «Академия», 2010. – 400 с. – 2500 экз. -ISBN 978-5-7695-6468-0.
2. Миронова Ю.Н. О -псевдокомпактных отображениях // Сибирский математический журнал. Май-июнь 2001. Том 42, №3, с. 634-644.
3. Миронова Ю.Н. О псевдокомпактных, счетно компактных, локально бикompактных отображениях и k-отображениях // Сибирский математический журнал. Сентябрь-октябрь 2002. Том 43, №5. Новосибирск, 2002, с. 1115-1129.
4. Миронова Ю.Н. Математические аспекты геоинформатики // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 7, №5 (2015) <http://naukovedenie.ru/PDF/93TVN515.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ. DOI: 10.15862/93TVN515.

5. Миронова Ю.Н. Трудности классификации геоинформационных объектов // Интернет-журнал «НАУКОВЕДЕНИЕ» Том 9, №1 (2017) <http://naukovedenie.ru/PDF/09TVN117.pdf> (доступ свободный). Загл. с экрана. Яз. рус., англ..
6. Миронова Ю.Н. Некоторые топологические свойства лексикографически упорядоченного квадрата // Международный журнал прикладных и фундаментальных исследований. – 2015. – № 12-10. – С. 1908-1909. URL: <http://www.applied-research.ru/ru/article/view?id=8408>
7. Яхьяева Г.Э. Нечеткие множества и нейронные сети: Учебное пособие. – М.: Интернет-Университет Информационных технологий; БИНОМ, 2006. – 316 с.
8. Mironova Yu.N. The classification of geoinformation objects // International Journal Of Applied And Fundamental Research. – 2016. – № 2 – URL: www.science-sd.com/464-25151 (29.09.2016).
9. Интернет ресурс «Геоинформатика. Под редакцией проф. В.С. Тикунова. М.: «Академия», 2005» <http://www.studfiles.ru/preview/1817795/>
10. Интернет ресурс «Орлов А.И. Прикладная статистика. М.: Издательство «Экзамен», 2004» <http://www.aup.ru/books/m163/>