

Романов Вадим Николаевич

ПРИМЕНЕНИЕ НЕЧЕТКИХ МОДЕЛЕЙ В КОРРЕЛЯЦИОННОМ АНАЛИЗЕ

В статье проведено исследование возможности применения нечетких моделей в корреляционном анализе и на их основе рассмотрена реализация метода главных компонент. Показаны преимущества предлагаемого подхода, позволяющего осуществлять совместную обработку количественных и качественных данных, уменьшить трудоемкость расчетов и нивелировать ошибки, связанные с несогласованностью данных.

Адрес статьи: www.gramota.net/materials/1/2013/12/40.html

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

Альманах современной науки и образования

Тамбов: Грамота, 2013. № 12 (79). С. 151-154. ISSN 1993-5552.

Адрес журнала: www.gramota.net/editions/1.html

Содержание данного номера журнала: www.gramota.net/materials/1/2013/12/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: almanac@gramota.net

Список литературы

1. Істоміна Г. В. Мистецтво народної кераміки Волині другої половини XIX – XX століть (типологія, стилістика, художні особливості): автореф. дисс. ... к. мистецтвознавства. Львів, 2007. 20 с.
2. Клименко О. О. Народна кераміка Опішні (до проблеми традицій та інновацій в народних художніх промислах): автореф. дисс. ... к. мистецтвознавства. Львів, 1995. 26 с.
3. Кравченко А. І. Українська музична культура в контексті дослідження проблем культурологічної регіоніки // Вісник Державної академії керівних кадрів культури і мистецтв: наук. журнал. К.: Міленіум, 2012. № 4. С. 139-143.
4. Кубицька Н. В. Гончарські осередки Кульчина та Рокити кінця XIX – XX ст. (історія, типологія, художні особливості): автореф. дисс. ... к. мистецтвознавства. Львів, 2007. 20 с.
5. Личковах В. А. Універсализм та регіоніка у сучасній етнокультурології // Дивосад культури: вибрані статті з естетики, культурології, філософії мистецтва. Чернівці: РВК «Деснянська правда», 2006. С. 120-126.
6. Мішанин В. Д. Гончарство північної групи малих осередків Опішненського гончарного району (друга половина XIX – XX століття): автореф. дисс. ... к.і.н. Львів, 2007. 20 с.
7. Нагорна Л. П. Регіональна ідентичність: український контекст. К.: ППІЕНД імені І. Ф. Кураса НАН України, 2008. 405 с.
8. Опішнянська мальована миска другої половини XIX – початку XX століття / автор-упорядник Олесь Пошивайло. Опішне: Українське народознавство, 2010. 640 с.
9. Петрушенко Р. М. Художні особливості розписів кераміки Бара XIX – поч. XX ст. [Електронний ресурс]. URL: <http://muzey.vn.ua/node/89.htm> (дата обращения: 31.10.2013).
10. Слободян О. О. Пістинський центр гончарства XIX – першої половини XX століття (історія, типологія, художні особливості, майстри): автореф. дисс. ... к. мистецтвознавства. Львів, 2001. 17 с.
11. Студенников І. Регіоналістика і регіонаознавство: проблеми методології та категоріального апарату // Регіональна історія України. К., 2007. Вип. 1. С. 67-78.
12. Шульгіна В. Д., Микуланінець Л. М. Регіональний контекст науково-дослідної діяльності кафедри теоретичної та прикладної культурології НАКККІМ // IX Культурологічні читання пам'яті Володимира Подкопаєва «Концептуальні проблеми розвитку української культури у світлі підготовки і проведення 2012 року як року культури та відродження музеїв в Україні»: збірник матеріалів міжнародної науково-практичної конференції (Київ, 2-3 червня 2011 р.). К.: НАКККІМ, 2011. Т. I. С. 243-245.

**COMPOSITIONAL FEATURES OF UKRAINIAN CERAMIC BOWLS DECORATION
(BY EXAMPLE OF POLTAVA, KIEV, PODOLYE REGIONS)**

Paslavskaya Lyudmila Aleksandrovna

*National Academy of Managing Staff of Culture and Arts, Ukraine
paslavska_ludmyla@i.ua*

The prospects of regional researches in culturology and art criticism are designated in the article. The Ukrainian ceramic bowls are considered in the context of local handicrafts. The genetic relationship of the composition model and decoration of ceramic bowls with national symbolics is shown. The comparative characteristics of three main composition models of ceramic bowls is presented by the example of the specimens of Poltava, Kiev, and Podolye regions of Ukraine.

Key words and phrases: regional studies in art criticism; arts and crafts; Ukrainian ceramic bowl; composition; decoration.

УДК 519.8

Физико-математические науки

В статье проведено исследование возможности применения нечетких моделей в корреляционном анализе и на их основе рассмотрена реализация метода главных компонент. Показаны преимущества предлагаемого подхода, позволяющего осуществлять совместную обработку количественных и качественных данных, уменьшит трудоемкость расчетов и нивелировать ошибки, связанные с несогласованностью данных.

Ключевые слова и фразы: корреляционный анализ; коэффициент корреляции; нечеткая модель; метод главных компонент.

Романов Вадим Николаевич, д.т.н., профессор

*Национальный минерально-сырьевой университет «Горный»
vromanvpi@mail.ru*

ПРИМЕНЕНИЕ НЕЧЕТКИХ МОДЕЛЕЙ В КОРРЕЛЯЦИОННОМ АНАЛИЗЕ[©]

Одним из применений корреляционного анализа является обнаружение «скрытых» параметров, лежащих в основе непосредственно наблюдаемых закономерностей. При изменении причины изменяются и связи между наблюдаемыми величинами. Корреляционный анализ характеризуется рядом ограничений. Количественные и

качественные данные обрабатываются раздельно, а затем проводится их согласование [1, с. 110]. Существенное значение имеет точность экспериментальных данных. Например, для определения рангового коэффициента корреляции Спирмена, применяемого при обработке качественных данных, с ошибкой 10% число объектов должно быть не менее 100, для обычного числового коэффициента корреляции объем данных еще больше. В реальных ситуациях приходится иметь дело с меньшим объемом данных, что увеличивает ошибку определения коэффициента корреляции. При расчете коэффициента корреляции величины должны иметь нормальное распределение [2, с. 384]. Следует также иметь в виду, что коэффициент корреляции является линейным приближением и учитывает только линейные взаимосвязи величин. В реальной ситуации при большом числе переменных взаимосвязи являются более сложными и могут включать нелинейные члены и взаимодействия. Кроме того, во многих случаях информация об объектах и критериях является неполной, разнородной, и имеется неопределенность в оценке объектов по критериям. Поэтому точные числовые расчеты не имеют смысла. Целью настоящей статьи является исследование возможности применения нечетких моделей в корреляционном анализе, что позволяет обобщить схему анализа на случай неопределенности и устранить отмеченные недостатки. Задача формулируется в следующем виде. Дано множество объектов $X = \{x_1, \dots, x_m\}$, которые оцениваются по множеству критериев (признаков, свойств) $K_1, \dots, K_n$, представленных нечеткими градациями в диапазоне [ОН, ОВ], где ОН – очень низкое значение, ОВ – очень высокое. Требуется определить корреляцию (взаимозависимость) между критериями и ее возможные причины. Алгоритм решения задачи включает следующие шаги:

1. Исходная количественная и качественная информация об объектах и критериях, полученная измерениями и экспертными методами, преобразуется в нечеткие градации способом, изложенным в [4, с. 123].
2. Определяется коэффициент корреляции для каждой пары критериев из известного соотношения, модифицированного для нашего случая

$$\rho_{lm} = \frac{\sum_{i=1}^n (K_i^l - \bar{K}^l)(K_i^m - \bar{K}^m)}{\left[\sum_{i=1}^n (K_i^l - \bar{K}^l)^2 \sum_{i=1}^n (K_i^m - \bar{K}^m)^2 \right]^{1/2}}, \quad (1)$$

с той только разницей, что все величины представлены в виде нечетких градаций, и вычисления выполняются с помощью правил нечеткой арифметики [3, с. 145]. В нашем случае нижний индекс в (1) соответствует объекту, а верхний – критерию.

3. Объекты распределяются по группам в соответствии со значениями коэффициента корреляции.
4. Для выявления «скрытых» параметров применяется метод главных компонент.
5. По результатам анализа делается заключение о возможных причинах наблюдаемых взаимосвязей.

Рассмотрим пример. Пусть исходные данные представлены в Табл. 1. Предполагается, что ошибка оценок в Табл. 1 составляет одну градацию.

Таблица 1. Исходные данные для примера

Объекты	Критерии						
	K^1	K^2	K^3	K^4	K^5	K^6	K^7
x_1	Н-С	С	С-В	Н-С	ОН-Н	ОН-Н	Н
x_2	С	С	Н-С	Н-С	В-ОВ	Н-С	С-В
x_3	С	Н-С	Н	ОН-Н	С-В	ОН-Н	ОН
x_4	С	Н	С	ОН-Н	С	ОН-Н	ОН
x_5	С	ОН-Н	Н-С	Н-С	Н-С	Н	ОН-Н
x_6	ОН-Н	ОН-Н	Н	Н	ОН-Н	Н	ОН
x_7	С	В	Н	Н-С	В-ОВ	Н-С	ОН
x_8	В	С	ОВ	В-ОВ	Н-С	ОВ	В
x_9	С-В	С	С	ОВ	ОВ	ОН-Н	ОН-Н
x_{10}	В-ОВ	Н	ОН	ОН-Н	ОВ	Н-С	ОН
x_{11}	С-В	Н	Н-С	ОН	ОВ	Н-С	ОН
x_{12}	Н-С	Н	Н-С	Н	С	ОН	ОН
x_{13}	В	ОН	ОН-Н	Н	В-ОВ	ОН	ОН
x_{14}	ОВ	С-В	С	Н-С	Н-С	С	С-В
x_{15}	С-В	В-ОВ	ОН-Н	Н	ОН	ОН-Н	ОН
x_{16}	С-В	ОН	Н	ОН-Н	Н-С	Н-С	ОН
x_{17}	ОН	ОН	ОН	Н	Н	Н	ОН
x_{18}	ОН	ОН-Н	С	ОН-Н	Н	ОН-Н	ОН-Н
x_{19}	С	С	Н	Н	В	Н	ОН-Н
x_{20}	В	ОВ	В	ОН	С-В	ОН-Н	ОН
$\bar{K}$	С	Н-С	Н-С	Н	С	Н	ОН-Н

Примечание. Здесь и в тексте использованы обозначения: ОН – очень низкое значение, Н – низкое, С – среднее, В – высокое, ОВ – очень высокое, (ОН-Н) – между очень низким и низким значениями, (Н-С) – между низким и средним, (С-В) – между средним и высоким, (В-ОВ) – между высоким и очень высоким.

На основе исходных данных по соотношению (1) рассчитывался коэффициент корреляции. Результаты расчетов приведены в Табл. 2.

Таблица 2. Матрица коэффициентов корреляции

	K^1	K^2	K^3	K^4	K^5	K^6	K^7
K^1	ООВ	Н-С	ОН-Н	ОН-Н	Н-С	Н-С	Н
K^2	Н-С	ООВ	Н-С	ОН-Н	ОН	ОН	ОН-Н
K^3	ОН-Н	Н-С	ООВ	Н-С	ОН ⁻	С	С
K^4	ОН-Н	ОН-Н	Н-С	ООВ	ОН	Н-С	С-В
K^5	Н-С	ОН	ОН ⁻	ОН	ООВ	ООН	ОН ⁻
K^6	Н-С	ОН	С	Н-С	ООН	ООВ	В
K^7	Н	ОН-Н	С	С-В	ОН ⁻	В	ООВ

Примечание. Здесь и далее использованы следующие обозначения: ООН – значение, соответствующее 0; ООВ – значение, соответствующее 1; знак «минус» сверху означает, что градация расположена слева от ООН (имеет отрицательное значение).

Для определения «скрытых» параметров используем метод главных компонент. Требуется привести корреляционную матрицу к диагональному виду, решив уравнение $AX = \lambda X$, где A – корреляционная матрица, λ – собственные числа, играющие роль скрытых параметров, X – собственные векторы. Решение осуществляется итеративно, причем все вычисления выполняются в нечетких градациях. Для вычисления первого собственного значения λ_1 сначала определяются суммы строк в Табл. 2, и наибольшее значение принимается за первое приближение для λ_1 . Затем все суммы делятся на наибольшее значение, и полученные нормированные значения принимаются в первом приближении за компоненты первого собственного вектора. Процесс повторяется, пока значение вектора не станет постоянным. Следует отметить, что для определения λ_1 требуется всего 4-5 итераций, т.е. меньше, чем при числовых расчетах. Это и понятно, так как используются сглаженные данные. Кроме того, случайные ошибки в пределах одной градации практически не влияют на результат. Расчеты дают $\lambda_1 = В/С^2$, что соответствует числовому значению 2,8. Собственному значению λ_1 соответствует собственный вектор, имеющий семь компонент $X_1 = ((С-В) С В ОН (В-ОВ) (В-ОВ))$. Матрица коэффициентов корреляции первого параметра (фактора) представлена в Табл. 3.

Таблица 3. Матрица коэффициентов корреляции первого параметра

X_1	С-В	С	В	В	ОН	В-ОВ	В-ОВ
С-В	Н-С	Н	Н-С	Н-С	ОН	С	С
С	Н	ОН-Н	Н-С	Н-С	ОН	Н-С	Н-С
В	Н-С	Н-С	С	С	ОН	С-В	С-В
В	Н-С	Н-С	С	С	ОН	С-В	С-В
ОН	ОН	ОН	ОН	ОН	ООН	ОН	ОН
В-ОВ	С	Н-С	С-В	С-В	ОН	С-В	С-В
В-ОВ	С	Н-С	С-В	С-В	ОН	С-В	С-В

Для исключения влияния первого параметра вычтем из матрицы в Табл. 2 матрицу, представленную в Табл. 3, что дает первую остаточную матрицу после исключения первого параметра. Результирующая матрица представлена в Табл. 4. По данным этой матрицы рассчитываются второе собственное значение λ_2 и соответствующий ему вектор X_2 . Вычисления выполняются аналогично расчетам для λ_1 и X_1 . Имеем $\lambda_2 = В/С$, что соответствует числовому значению 1,4. Собственный вектор X_2 , соответствующий λ_2 , имеет вид $X_2 = ((С-В) (Н-С) Н⁻ ОН⁻ (В-ОВ) (ОН-Н)⁻ Н⁻)$.

Для его вычисления достаточно 5-6 итераций.

Таблица 4. Остаточная матрица после исключения первого параметра

	K^1	K^2	K^3	K^4	K^5	K^6	K^7
K^1	С-В	ОН	(ОН-Н) ⁻	(ОН-Н) ⁻	Н	ОН ⁻	(ОН-Н) ⁻
K^2	ОН	В-ОВ	ООН	(ОН-Н) ⁻	ООН	Н ⁻	(ОН-Н) ⁻
K^3	(ОН-Н) ⁻	ООН	С	ОН ⁻	(ОН-Н) ⁻	ОН ⁻	ОН ⁻
K^4	(ОН-Н) ⁻	(ОН-Н) ⁻	ОН ⁻	С	ООН	(ОН-Н) ⁻	ООН
K^5	Н	ООН	(ОН-Н) ⁻	ООН	ООВ	ОН ⁻	(ОН-Н) ⁻
K^6	ОН ⁻	Н ⁻	ОН ⁻	(ОН-Н) ⁻	ОН ⁻	Н-С	ОН
K^7	(ОН-Н) ⁻	(ОН-Н) ⁻	ОН ⁻	ООН	(ОН-Н) ⁻	ОН	Н-С

Дальнейшие вычисления выполняются аналогично предыдущему. Корреляционная матрица коэффициентов второго параметра (фактора) представлена в Табл. 5.

Таблица 5. Матрица коэффициентов корреляции второго параметра

X_2	C-B	H-C	H ⁻	OH ⁻	B-OB	(OH-H) ⁻	H ⁻
C-B	H-C	OH-H	(OH-H) ⁻	OH ⁻	C	OH ⁻	(OH-H) ⁻
H-C	OH-H	OH-H	OH ⁻	OOH	H	OH ⁻	OH ⁻
H ⁻	(OH-H) ⁻	OH ⁻	OH	OOH	(OH-H) ⁻	OH	OH
OH ⁻	OH ⁻	OOH	OOH	OOH	OH ⁻	OOH	OOH
B-OB	C	H	(OH-H) ⁻	OH ⁻	C-B	(OH-H) ⁻	(OH-H) ⁻
(OH-H) ⁻	OH ⁻	OH ⁻	OH	OOH	(OH-H) ⁻	OOH	OH
H ⁻	(OH-H) ⁻	OH ⁻	OH	OOH	(OH-H) ⁻	OH	OH

Для исключения второго параметра вычтем из матрицы в Табл. 4 матрицу, представленную в Табл. 5, что дает остаточную матрицу после исключения второго параметра. Результирующая матрица представлена в Табл. 6.

Таблица 6. Остаточная матрица после исключения второго параметра

	K^1	K^2	K^3	K^4	K^5	K^6	K^7
K^1	OH-H	OH ⁻	OOH	OH ⁻	(OH-H) ⁻	OOH	OOH
K^2	OH ⁻	C-B	OH	(OH-H) ⁻	H ⁻	(OH-H) ⁻	OH ⁻
K^3	OOH	OH	H-C	OH ⁻	OOH	(OH-H) ⁻	(OH-H) ⁻
K^4	OH ⁻	(OH-H) ⁻	OH ⁻	C	OH	(OH-H) ⁻	OOH
K^5	(OH-H) ⁻	H ⁻	OOH	OH	H-C	OH	OOH
K^6	OOH	(OH-H) ⁻	(OH-H) ⁻	(OH-H) ⁻	OH	H-C	OOH
K^7	OOH	OH ⁻	(OH-H) ⁻	OOH	OOH	OOH	H

Из Табл. 6 видно, что сумма элементов в строках этой матрицы составляет одну-две градации, т.е. находится в пределах ошибки, кроме того ни одно из значений матрицы не превышает градацию В (высокое значение), т.е. строгое условие достоверности не выполняется. Поэтому дальнейшие вычисления являются ненадежными, и достаточно ограничиться двумя собственными значениями, вклад которых составляет $(B/C^2 + B/C)/(B/OH) = 0,6$. Ясно, что ошибка оценок в матрице возрастает с увеличением номера параметра. Она особенно сказывается на точности определения собственных векторов, начиная с X_3 ; на точности расчета собственных чисел ошибка сказывается меньше, так как они определяются максимальным значением суммы в строках соответствующей матрицы. Поэтому мы ограничились расчетом собственных значений, чтобы оценить порядок величин. Вычисления дают $\lambda_3 = C/C = OOB$, $\lambda_4 = B$, $\lambda_5 = H-C$. Последнее значение находится за порогом достоверности, поэтому расчеты λ_6 и λ_7 не имеют смысла. Таким образом, из проведенного рассмотрения следует, что надежно определяются два первых собственных значения и соответствующие им собственные векторы. Тем самым устанавливается объективный критерий прекращения расчетов. Следует отметить, что при использовании нечетких градаций все результаты получаются при меньшем числе итераций, чем при числовом представлении данных, являются менее трудоемкими и могут быть выполнены вручную. Что касается интерпретации причин взаимосвязей, то она определяется предметной областью. Например, при анализе качества обучения скрытыми параметрами могут быть способности и интерес к учебе; при анализе качества продукции роль скрытых параметров играют уровень технологии и качество сырья; при диагностировании заболеваний – наследственность и системные изменения в организме и т.п.

Список литературы

1. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989. 540 с.
2. Кендалл М., Стьюарт А. Статистические выводы и связи. М.: Наука, 1973. 899 с.
3. Романов В. Н. Нечеткие модели принятия решений // Альманах современной науки и образования. Тамбов: Грамота, 2013. № 5 (72). С. 144-147.
4. Романов В. Н. Определение существенных признаков в задачах идентификации топологическими методами // Альманах современной науки и образования. Тамбов: Грамота, 2013. № 7 (74). С. 122-129.

USING FUZZY MODELS FOR CORRELATION ANALYSIS

Romanov Vadim Nikolaevich, Doctor in Technical Sciences, Professor
National Mineral Resource University (University of Mines)
vromanvpi@mail.ru

In the article the potential research of using fuzzy models for correlation analysis is made, on this basis the realization of the method of principal components is considered. The advantages of the suggested approach, which allows making joined quantity and quality data processing, reducing the laboriousness of calculations and leveling mistakes connected with data mismatching are shown.

Key words and phrases: correlation analysis; correlation coefficient; fuzzy model; method of principal components.