УДК 519.237.7

М.Е. Королёв, канд. физ-мат. наук, Н.Н. Дудникова, канд. техн. наук,
Е.А. Королёв, канд. техн. наук, О.Н. Куктенко

Автомобильно-дорожный институт
ГОУВПО «Донецкий национальный технический университет», г. Горловка

ИСПОЛЬЗОВАНИЕ МНОГОМЕРНЫХ СТАТИСТИЧЕСКИХ МЕТОДОВ ФАКТОРНОГО АНАЛИЗА В ИССЛЕДОВАНИИ АВТОМОБИЛЬНЫХ РЫНКОВ

Проанализирована и доказана экспериментально классическая модель факторного анализа на примере статистических данных автомобильных рынков. Выделены латентные факторы, построена факторная плоскость с нанесением объекта исследования – автомобильных рынков.

Ключевые слова: *анализ статистический, многомерное шкалирование, метод главных факторов, математическая модель*

Введение

В настоящее время работа по освоению методов многомерного статистического анализа (МСА) и внедрению их в аналитическую практику становится особенно актуальной, поскольку отслеживание и адекватная реакция на различные изменения в любой сфере человеческой деятельности возможны при наличии системы объективного отражения и оперативной передачи достаточно полной микро- и макроуровневой информации.

В современных условиях особенно значимым становится изучение комплексов показателей, представляющих различные сферы исследований. Многомерные методы длительное время широко использовались в практической статистике стран Европы, Америки, Азии, где созданы и функционируют технологичные системы обработки и передачи многомерных данных в компьютерных сетях [1–8]. Знания в области МСА представляются необходимыми для овладения новейшими методологическими разработками в области теоретической статистики и ее развития.

МСА основывается на теоретической базе высшей математики и математической статистики. Множество его методов разбивается на две большие группы. К первой относятся методы, применение которых предполагает знание законов распределения многомерной случайной величины и позволяют производить статистическую оценку явлений и процессов, проверять статистические гипотезы – это методы вероятностного анализа многомерных данных. Ко второй группе принадлежат методы, для применения которых не обязательно знание законов распределения, но существенна рациональная логическая конструкция, позволяющая адекватно моделировать реальные процессы и явления – это методы логико-алгебро-геометрического направления [9]. Совокупность методов, которые относятся к двум названным выше направлениям МСА, позволяют решить разнообразные задачи и область их применения четко не разграничена. Объекты и явления рассматриваются с учетом одновременно некоторого множества признаков. Это позволяет добиваться в исследованиях полноты теоретического описания наблюдаемых объектов и объективности последующих выводов.

Реально изучаемые объекты и явления имеют практически всегда многопризнаковую природу, надежное отображение их в многообразных математических моделях возможно при условии учета комплекса присущих им наиболее существенных характеристик.

МСА – это совокупность глубоко формализованных статистических методов, базирующихся на представлении исходной информации в многомерном геометрическом пространстве и позволяющих определять неявные (латентные), но объективно существующие закономерности в организационной структуре и тенденциях развития изучаемых социально-экономических явлений и процессов [11–12].

Применение методов МСА требует творческого подхода к решению аналитических задач, поскольку они весьма разнообразны и многочисленны. Для решения даже одного типа задач здесь существуют десятки и сотни различных приемов из кластер-анализа, факторного анализа, многомерного шкалирования и т. д. Чтобы правильно выбрать тот или иной метод или комплекс методов для последовательного решения поставленной проблемы, необходимы знания, профессионализм и интуиция. Это становится особенно важным при интерпретации аналитических результатов, часто неоднозначных, когда выводы должны отвечать логической схеме сложных математических расчетов. Противоречия в выводах свидетельствуют о некорректности решения задачи или некорректности интерпретации аналитических результатов.

В МСА обрабатываются многомерные (многопризнаковые) совокупности данных. Число признаков (размерность совокупности) при этом может быть – от 1 до 100 и более и в конечном счете определяется условиями задачи и целями исследования [9].

Практическое применение методов МСА требует обязательного использования вычислительной техники. Можно сказать, что эти методы в силу сложности и трудоемкости нереализуемы без технических средств.

Таким образом, методы многомерного статистического анализа сегодня называют интеллектуальным инструментом исследователя и активно применяются в аналитической практике.

Большой интерес вызывают методы многомерного шкалирования (МШ), разработанные и применяемые в практике для исследований сложных явлений и процессов, не поддающихся непосредственному описанию или моделированию. В основу теории многомерного шкалирования положена идея о возможности развертывания наблюдаемых объектов в некотором теоретическом пространстве, адекватно отображающем реальность [12].

В отличие от других статистических методов поиск координатного пространства в МШ осуществляется не по значениям самих характеризующих объекты признаков, а по данным, представляющим различия, или наоборот, сходство этих объектов. Основным источником данных здесь являются в одних случаях эксперты, субъективно воспринимающие и оценивающие относительное расположение объектов наблюдения в реальных условиях, в других – результаты прямой регистрации сведений о состоянии и поведении объектов.

Цель аналитической работы с данными – определение местонахождения объекта в «пространстве восприятия (субъектов)» и создание его образа. Имеется в виду, что непосредственно о самом объекте даже по значениям некоторого набора признаков нельзя судить достаточно надежно или полно. В то же время эксперты или просто наблюдатели еще до проведения аналитических расчетов видят, интуитивно чувствуют различия изучаемых объектов. Неосознанные, нечеткие представления об объектах должны быть конкретизированы и это осуществимо в теоретическом «пространстве восприятия», построенном по субъективным оценкам. В этом представляемом пространстве проявляют себя латентные факторы, становится очевидным действие этих факторов на пространственное расположение объекта, измеримо расстояние между объектами [9].

Цель

Необходимо определить местонахождение объекта «автомобильный рынок» в «пространстве восприятия субъектов-потребителей», создать его образ – визуализацию на Евклидовой плоскости и найти латентные факторы.

Основной материал исследования

Широко применяется также факторный анализ [10], представляющий собой совокупность методов, которые на основе реально существующих связей признаков (или объектов) позволяют выявлять латентные (скрытые) обобщающие характеристики организационной структуры и механизма развития изучаемых явлений и процессов.

Из числа методов, позволяющих обобщать значения элементарных признаков, выделяется метод главных факторов, который отличается своей простой логической конструкцией. Он дает возможность по m – числу исходных признаков выделить m главных факторов, или обобщенных признаков. Математическая модель главных факторов базируется на логичном допущении, что значения множества взаимосвязанных признаков порождают некоторый общий результат.

Учитывая вышесказанное, для решения нижеприведенной задачи применим следующую последовательность методов анализа:

1) метод многомерного шкалирования Торгерсона [9], основанный на использовании количественных характеристик объектов;

2) метод главных факторов [9] (метод Хоттеллинга), поскольку он позволяет сравнительно быстро выделить наибольшее количество общих факторов, учитывающих почти всю суммарную общность.

Пусть экспертами, например профессиональными экономистами, оценивались четыре автомобильных рынка (A, B, C, D), расположенных в разных городах и имеющих различные условия торговли. При этом: A(X1) – рынок с определенно неблагоприятными условиями для покупателя, где к продаже представлено небольшое количество автомобилей, сервисные услуги ограничены и оказываются неквалифицированными работниками, имеется жесткая схема оплаты за покупку, например только наличным расчетом; B(X2) – с широким выбором автомобилей (по маркам, уровню цен и т. д.), разнообразными сервисными услугами, в целом наиболее благоприятными условиями для покупателя; C(X3) – с ограниченным выбором автомобилей, суженным кругом услуг, незначительными расхождениями в ценах на автомобили, то есть ограниченными возможностями для покупателя; D(X4) – рынок, специализированный на автомобилях одной марки, но предоставляющий разнообразные услуги, в общем, имееющий благоприятные условия для покупателя.

Свои «оценки» эксперты должны были дать по ранее сложившимся представлениям после ознакомления с соответствующей информацией и общения с покупателями, в результате чего изучаемый рынок следовало отнести к классу A, B, C или D.

Необходимо определить местонахождение объекта «автомобильный рынок» в «пространстве восприятия субъектов-потребителей», создать его образ – визуализацию на Евклидовой плоскости и найти латентные факторы.

Применим классическую модель многомерного шкалирования: дана матрица идентификаций, строки которой – рынки автомобилей, столбцы – характерные признаки.

Для решения поставленной задачи необходимо выполнить следующие этапы: стандартизировать исходные данные (таблица 1); определить матрицы различий по стандартизированным данным; применить классическую модель многомерного шкалирования.

Таблица 1 – Исходные данные

	Признак		
Объект	ЭКСПЕРТ 1 = Р 1	ЭКСПЕРТ 2 = Р 2	ЭКСПЕРТ 3 = Р 3
А = Х1	28,8	446,6	80,4
В = Х2	18,5	105,8	27,1
С = Х3	19	141,6	28,5
Д = Х4	20,1	2,9	43,8
Среднее значение	21,6	174,225	44,95

В свою очередь применение модели включает: вычисление «средних» характеристик; преобразование матрицы различий в матрицу с двойным центрированием.

Итак, были стандартизированы исходные данные $z_{ij} = \nu_{ij} / \overline{\nu_j}$.

$$z_{ij} = \begin{pmatrix} 1{,}33 & 2{,}56 & 1{,}79 \\ 0{,}86 & 0{,}61 & 0{,}60 \\ 0{,}88 & 0{,}81 & 0{,}63 \\ 0{,}93 & 0{,}02 & 0{,}97 \end{pmatrix}.$$

Построена матрица мер различия профилей, причем из большого числа метрических формул использовано евклидово расстояние: $\delta_{ij}^E = \left(\sum_k (\nu_{ik} - \nu_{jk})^2 \right)^{1/2}$.

$$\Delta = \begin{pmatrix} 0 & 2{,}337 & 2{,}146 & 2{,}704 \\ 2{,}337 & 0 & 0{,}209 & 0{,}702 \\ 2{,}146 & 0{,}209 & 0 & 0{,}867 \\ 2{,}704 & 0{,}702 & 0{,}867 & 0 \end{pmatrix}.$$

Значения элементов матрицы с двойным центрированием $\Delta^* = (\delta_{ij}^*)$ найдены по формуле [6]:

$$\delta_{ij}^* = -\tfrac{1}{2}(\delta_{ij}^2 - \delta_{i\cdot}^2 - \delta_{\cdot j}^2 + \delta_{\cdot\cdot}^2),$$

где $\delta_{i\cdot}^2 = \frac{1}{j}\sum_j \delta_{ij}^2$ – средняя для характеристик различий в j-х столбцах i-й строки, возведенных в квадрат;

$\delta_{\cdot j}^2 = \frac{1}{i}\sum_i \delta_{ij}^2$ – средняя для характеристик различий в i-х строках j-го столбца, возведенных в квадрат; $\delta_{\cdot\cdot}^2 = \frac{1}{ij}\sum_i \sum_j \delta_{ij}^2$ – средняя величина для квадратов характеристик различий матрицы Δ.

Матрица с двойным центрированием имеет вид

$$\Delta^* = \begin{pmatrix} 3{,}177 & -0{,}975 & -0{,}621 & -1{,}581 \\ -0{,}975 & 0{,}333 & 0{,}236 & 0{,}406 \\ -0{,}621 & 0{,}236 & 0{,}183 & 0{,}202 \\ -1{,}581 & 0{,}406 & 0{,}202 & 0{,}972 \end{pmatrix}.$$

Для матрицы с двойным центрированием применена классическая модель главных факторов с корреляционной матрицей $R_h = \Delta^*$ (таблица 2)

$$R_h = \Delta^* = \begin{pmatrix} 3{,}177 & -0{,}975 & -0{,}621 & -1{,}581 \\ -0{,}975 & 0{,}333 & 0{,}236 & 0{,}406 \\ -0{,}621 & 0{,}236 & 0{,}183 & 0{,}202 \\ -1{,}581 & 0{,}406 & 0{,}202 & 0{,}972 \end{pmatrix}.$$

Для дальнейших расчетов был использован метод главных факторов – алгоритм Хоттелинга.

Таблица 2 – Корреляционная матрица

Признак	R_h				$S_i^{(1)} = \sum_j r_{ij}$	$\alpha^{(1)} = \frac{S_i^{(1)}}{S_{\max}}$
	X1	X2	X3	X4		
X1	3,177	–0,975	–0,622	–1,581	8,9E–16	0,667
X2	–0,975	0,333	0,2362	0,406	1,1E–15	0,833
X3	–0,622	0,236	0,183	0,202	1,3E–15	1
X4	–1,581	0,406	0,203	0,972	1,3E–15	1

Указанный алгоритм реализуется при выполнении нескольких итераций, каждая из которых содержит определенное количество циклов.

Первая итерация приведена в таблице 3.

Первый цикл итерации – это возведение в квадрат корреляционной матрицы.

Таблица 3 – Первый цикл итерации

Признак	$R_h^2 = R_h^{/} R_h$				$S_i^2 = \sum_j r_{ij}$	$P_i^2 = R_h S_i^1$	$\alpha_i^2 = \frac{P_i^2}{P_{\max}}$	$d_i = \left\|\alpha_i^2 - \alpha_i^1\right\|$
	X1	X2	X3	X4				
X1	13,9297	–4,211	–2,6379	–7,0807	0	–1E–15	–1,000	1,667
X2	–4,211	1,28209	0,80984	2,11908	8,9E–16	3,6E–16	0,301	0,533
X3	–2,6379	0,80984	0,51643	1,31166	2,2E–16	2,2E–16	0,187	0,813
X4	–7,0807	2,11908	1,31166	3,65001	1,3E–15	6,1E–16	0,512	0,488

Второй цикл итерации – это возведение корреляционной матрицы в четвертую степень (таблица 4).

Таблица 4 – Второй цикл итерации

Признак	$R_h^4 = R_h^2 R_h^2$				$S_i^3 = \sum_j r_{ij}$	$P_i^3 = R_h^2 S_i^2$	$\alpha_i^3 = \frac{P_i^3}{P_{\max}}$	$d_i = \left\|\alpha_i^3 - \alpha_i^2\right\|$
	X1	X2	X3	X4				
X1	268,864	–81,198	–50,806	–136,86	0	–1E–14	–1,000	0,000
X2	–81,198	24,5227	15,3444	41,3309	–7E–15	4,1E–15	0,301	0,000
X3	–50,806	15,3444	9,60164	25,8595	3,6E–15	2,6E–15	0,188	0,000
X4	–136,86	41,3309	25,8595	69,6705	1,4E–14	7E–15	0,511	0,001

После возведения корреляционной матрицы в квадрат и в четвертую степень видно, что разности d_i резко уменьшаются, что говорит о корректности проводимых расчетов.

Третий цикл итерации – это возведение корреляционной матрицы в восьмую степень (таблица 5).

Таблица 5 – Третий цикл итерации

Признак	$R_h^8 = R_h^4 R_h^4$				$S_i^4 = \sum_j r_{ij}$	$P_i^4 = R_h^3 S_i^3$	$\alpha_i^4 = \frac{P_i^4}{P_{\max}}$	$d_i = \left\|\alpha_i^4 - \alpha_i^3\right\|$
	X1	X2	X3	X4				
X1	100193	–30259	–18933	–51002	0	–2E–12	–1,000	0,000
X2	–30259	9138,16	5717,72	15402,7	–4E–12	4,7E–13	0,302	0,001
X3	–18933	5717,72	3577,56	9637,42	0	2,9E–13	0,189	0,001
X4	–51002	15402,7	9637,42	25961,8	7,3E–12	7,9E–13	0,509	0,002

Оценки *S* и *P* подтверждают правильность проведенных вычислений, поэтому оценки компонента первого собственного вектора можно считать достоверными.

Четвертый цикл итерации – это возведение корреляционной матрицы в шестнадцатую степень (таблица 6).

Таблица 6 – Четвертый цикл итерации

Признак	$R_h^{16} = R_h^8 R_h^8$				$S_i^5 = \sum_j r_{ij}$	$P_i^5 = R_h^4 S_i^4$	$\alpha_i^5 = \dfrac{P_i^5}{P_{\max}}$	$d_i = \left\lvert \alpha_i^5 - \alpha_i^4 \right\rvert$
	X1	X2	X3	X4				
X1	1,4E+10	–4E+09	–3E+09	–7E+09	–1E–06	–3E–07	–1,000	0,000
X2	–4E+09	1,3E+09	7,9E+08	2,1E+09	2,4E–07	7,9E–08	0,302	0,000
X3	–3E+09	7,9E+08	5E+08	1,3E+09	–5E–07	4,9E–08	0,189	0,000
X4	–7E+09	2,1E+09	1,3E+09	3,6E+09	–1E–06	1,3E–07	0,509	0,000

После проведения четвертого цикла первая итерация была прервана, так как разности d_i близки к нулю.

Пятый цикл итерации – это возведение корреляционной матрицы в тридцать вторую степень. Он проводится для уточнения проведенных расчетов (таблица 7).

Таблица 7 – Пятый цикл итерации

Признак	$R_h^{32} = R_h^{16} R_h^{16}$				$S_i^6 = \sum_j r_{ij}$	$P_i^6 = R_h^5 S_i^5$	$\alpha_i^6 = \dfrac{P_i^6}{P_{\max}}$	$d_i = \left\lvert \alpha_i^6 - \alpha_i^5 \right\rvert$
	X1	X2	X3	X4				
X1	2,7E+20	–8E+19	–5E+19	–1E+20	–16384	–6262,9	–1,000	0,000
X2	–8E+19	2,4E+19	1,5E+19	4,1E+19	8192	1891,41	0,302	0,000
X3	–5E+19	1,5E+19	9,6E+18	2,6E+19	8192	1183,45	0,189	0,000
X4	–1E+20	4,1E+19	2,6E+19	7E+19	8192	3188,05	0,509	0,000

Таким образом, собственный вектор – это ненормированные значения α_i^5.

Далее были определены нагрузки первого главного фактора (таблица 8).

Таблица 8 – Вычисление нагрузок первого главного фактора

Признак	$\alpha_i^5 = U_1$	$\beta_1 = R_h \cdot \alpha_i^5$	$A = \dfrac{U_1 \sqrt{\lambda_1}}{\sum U_{1i}^2 \;\; ^{1/2}}$
X1	–1,000	–4,394	–1,780
X2	0,302	1,327	0,538
X3	0,189	0,830	0,336
X4	0,509	2,237	0,906

Первое собственное число $\max \beta_1 = \lambda_1 = 4{,}394$, вектор факторных нагрузок A = (–1,780 0,538 0,336 0,906).

Для определения необходимости проведения второй итерации с поиском второго собственного числа и вектора факторных нагрузок была найдена матрица парной корреляции $R_h^+ = A \cdot A'$, которая имеет вид:

$$R_h^+ = \begin{pmatrix} -1{,}780 \\ 0{,}538 \\ 0{,}336 \\ 0{,}906 \end{pmatrix} \cdot \begin{pmatrix} 1{,}780 & 0{,}538 & 0{,}336 & 0{,}906 \end{pmatrix} = \begin{pmatrix} 3{,}170 & -0{,}957 & -0{,}599 & -1{,}614 \\ -0{,}957 & 0{,}289 & 0{,}181 & 0{,}487 \\ -0{,}599 & 0{,}181 & 0{,}113 & 0{,}305 \\ -1{,}614 & 0{,}487 & 0{,}305 & 0{,}821 \end{pmatrix}.$$

Разность матриц $R_h - R_h^+$ покажет остаточную, не объясненную первым главным фактором, корреляцию и поможет ответить на вопрос о целесообразности выделения второго главного фактора.

$$R_1 = R_h - R_h^+ = \begin{pmatrix} 3{,}177 & -0{,}975 & -0{,}621 & -1{,}581 \\ -0{,}975 & 0{,}333 & 0{,}236 & 0{,}406 \\ -0{,}621 & 0{,}236 & 0{,}183 & 0{,}202 \\ -1{,}581 & 0{,}406 & 0{,}202 & 0{,}972 \end{pmatrix} - \begin{pmatrix} 7{,}105 & -2{,}614 & -1{,}979 & -2{,}511 \\ -2{,}614 & 0{,}962 & 0{,}728 & 0{,}924 \\ -1{,}979 & 0{,}728 & 0{,}551 & 0{,}700 \\ -2{,}511 & 0{,}924 & 0{,}700 & 0{,}888 \end{pmatrix} =$$

$$= \begin{pmatrix} 0{,}007 & -0{,}018 & -0{,}022 & 0{,}033 \\ -0{,}018 & 0{,}043 & 0{,}055 & -0{,}081 \\ -0{,}022 & 0{,}055 & 0{,}070 & -0{,}103 \\ 0{,}033 & -0{,}081 & -0{,}103 & 0{,}15 \end{pmatrix}.$$

Как видно, матрица первых остаточных коэффициентов корреляции содержит еще достаточно большие величины и вполне допускает оценку второго главного фактора. Последующее выполнение второй итерации аналогично первой, но вычисления производятся на данных матрицы остатков R_1.

В результате проведения двух итераций получены два главных фактора F_1 и F_2 (таблица 9).

Таблица 9 – Итоговая таблица

Главный фактор (факторные нагрузки)		
	F_1	F_2
A–X1	–1,780	0,085
B–X2	0,538	–0,208
C–X3	0,336	–0,265
D–X4	0,906	0,389

Далее было проведено графическое отображение главного фактора (рисунок 1) и его логическая интерпретация.

Проверка произведенных вычислений заключается в том, чтобы для известной матрицы скалярных произведений было верно равенство $\Delta^* = XX'$, т. е.:

$$XX' = \begin{pmatrix} -1{,}780 & 0{,}085 \\ 0{,}538 & -0{,}208 \\ 0{,}336 & -0{,}265 \\ 0{,}906 & 0{,}389 \end{pmatrix} \times \begin{pmatrix} -1{,}780 & 0{,}538 & 0{,}336 & 0{,}906 \\ 0{,}085 & -0{,}208 & -0{,}265 & 0{,}389 \end{pmatrix} = \begin{pmatrix} 3{,}177 & -0{,}975 & -0{,}621 & -1{,}581 \\ -0{,}975 & 0{,}333 & 0{,}236 & 0{,}406 \\ -0{,}621 & 0{,}236 & 0{,}183 & 0{,}202 \\ -1{,}581 & 0{,}406 & 0{,}202 & 0{,}972 \end{pmatrix}.$$

Последняя матрица совпадает с матрицей Δ^* – матрицей с двойным центрированием, что подтверждает правильность практического результата.

На рисунке 1 представлено факторное пространство – местонахождение объекта «автомобильный рынок» в «пространстве восприятия субъектов–потребителей», где: A – неблагоприятные условия; B – широкий выбор автомобилей, разнообразные сервисные услуги; C – ограниченный выбор автомобилей, суженный круг услуг; D – специализация на автомобилях одной марки, предоставление разнообразных услуг.

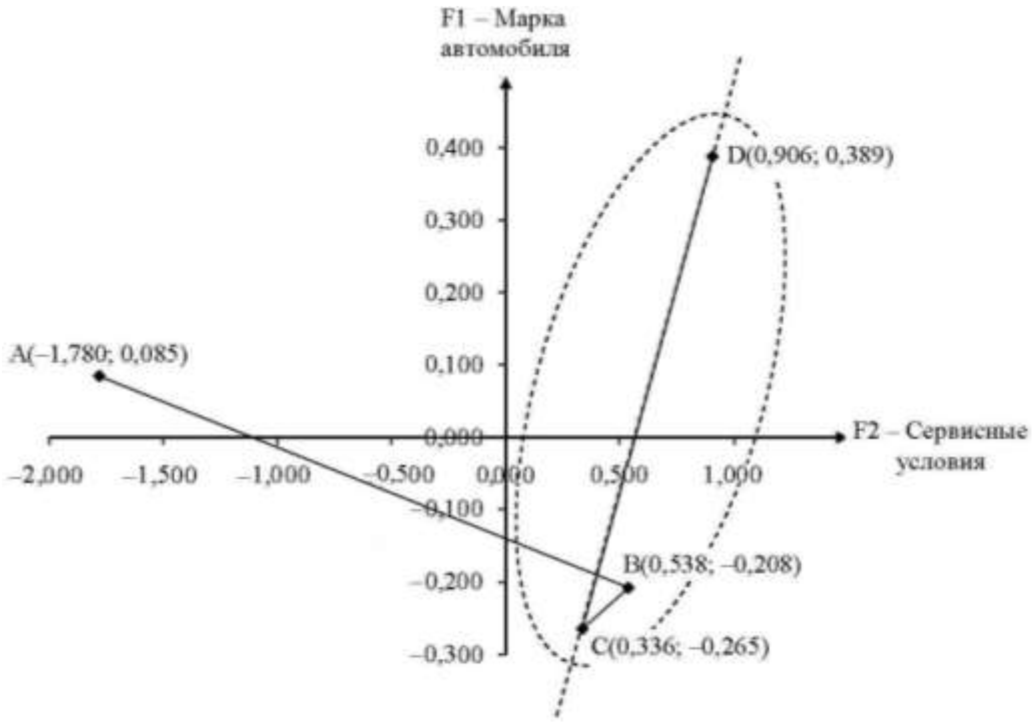

Рисунок 1 – Двумерная конфигурация шкал

Выводы

В данной работе были получены и доказаны экспериментально латентные факторы на основе статистических данных автомобильных рынков, действие которых на изучаемый объект скрыто от исследователя. Проанализирована и протестирована автоматизированная классическая модель факторного анализа; произведена визуализация и интерпретация результатов; выполнена математическая проверка самого метода. Полученные результаты позволяют считать успешной выполненную автоматизацию модели с исходной корреляционной матрицей. Это позволит использовать ее на практике при экономическом анализе, где количество рассматриваемых объектов существенно велико.

Список литературы

1. Глинский В.В., Ионин В.Г. Статистический анализ. М. – Новосибирск: ИНФРА-М, Сибирское соглашение, 2009. 241 с.
2. Дронов С.В. Многомерный статистический анализ. Барнаул: Изд-во Алтайского гос. университета, 2011. 239 с.
3. Дубров А.М., Мхитарян В.С. Многомерные статистические методы и основы эконометрики. М.: ИНФРА-М, 2007. 360 с.
4. Симчера В.М. Методы многомерного анализа статистических данных. М.: Финансы и статистика, 2008. 400 с.
5. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. М: Финансы и статистика, 2008. 352 с.
6. Borg I., Groenen P. J. F Modern Multidimensional Scaling: Theory and Applications. Springer, 2013. 636 p.
7. Borg I., Groenen P. J. F., Mair P. Applied Multidimensional Scaling. Springer, 2012. 122 p.
8. Härdle W. K., Simar L. Applied Multivariate Statistical Analysis. 4th Edition. Springer, 2015. 581 p.
9. Многомерный статистический анализ в экономике / Л.А. Сошникова [и др.]. М.: ЮНИТИ-ДАНА, 2012. 598 с.

10. Королёв М.Е., Королёв Е.А., Гетьманская В.Л. Компьютерные технологии прикладной математики в управлении предприятием // Инновационные перспективы Донбасса [Электронный ресурс]: материалы международной научно-практической конференции (г. Донецк, 20-22 мая 2015 г.) Т. 6: Актуальные проблемы инновационного развития экономики Донбасса; редкол.: Л.П. Полякова и др. / М-во образования и науки ДНР. Донецк: ГВУЗ «ДонНТУ», 2015. С. 182–186.
11. Горкавий В.К., Ярова В.В. Математична статистика. К.: Професіонал, 2009. 384 с.
12. Терещенко О.В. Многомерный статистический анализ данных в социальных науках. Минск: БГУ, 2012. 239 с.

М.Е. Королёв, Н.Н. Дудникова, Е.А. Королёв, О.Н. Куктенко
Автомобильно-дорожный институт
ГОУВПО «Донецкий национальный технический университет», г. Горловка

Использование многомерных статистических методов факторного анализа в исследовании автомобильных рынков

В данной работе проанализирована и протестирована автоматизированная классическая модель факторного анализа на основе статистических данных автомобильных рынков. Использован метод многомерного шкалирования Торгерсона, основанный на использовании количественных характеристик объектов и метод главных факторов (метод Хоттеллинга), поскольку он позволяет сравнительно быстро выделить наибольшее количество общих факторов, учитывающих почти всю суммарную общность. Произведена визуализация и интерпретация результатов, а именно определено местонахождение объекта «автомобильный рынок» в «пространстве восприятия субъектов-потребителей», создан его образ – визуализация на Евклидовой плоскости. Найдены и доказаны экспериментально латентные факторы, действие которых на изучаемый объект скрыто от исследователя. Выполнена математическая проверка использованного метода.

Полученные результаты позволяют сделать вывод о том, что автоматизацию модели с исходной корреляционной матрицей можно считать успешной. Это позволит использовать ее на практике при экономическом анализе, где количество рассматриваемых объектов существенно велико.

АНАЛИЗ СТАТИСТИЧЕСКИЙ, МНОГОМЕРНОЕ ШКАЛИРОВАНИЕ, МЕТОД ГЛАВНЫХ ФАКТОРОВ, МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

M.E. Korolev, N.N. Dudnikova, E.A. Korolev, O.N. Kuktenko
Automobile and Highway Institute of Donetsk National Technical University, Gorlovka

Multidimensional Statistical Methods of the Factor Analysis in the Car Market Research

In the work automated classical model of the factor analysis on the basis of statistical data of the car market is analyzed and tested. Torgerson’s multidimensional scaling method based on the object quantitative characteristics and principal-factor analysis (Hotelling method) is used. It allows to factorize comparatively quickly the greatest number of general factors taking into account almost all total generality. Visualization and interpretation of results are done, namely an object – car market in perception space (subjects-customers) is located, and its image is created – visualization on Euclidean plane. Experimental and latent factors which effect on the body of interest is concealed from researcher are found and proved. Mathematical examination of the used method is carried out.

Obtained results allow to make a conclusion that model automation with initial correlation matrix may be considered effective. It allows to use it in practice at the experimental analysis where number of considered objects is substantially large.

STATISTICAL ANALYSIS, MULTIDIMENSIONAL SCALING, PRINCIPAL-FACTOR ANALYSIS, MATHEMATICAL MODEL

Сведения об авторах

М.Е. Королёв
SPIN-код: 4980-9607
Телефон: 0505385168
Эл. почта: kustokust@gmail.com

Н.Н. Дудникова
SPIN-код: 1424-1363
Телефон: 050 589 90 37
Эл. почта: DudnikovaNN@rambler.ru

Е.А. Королёв
Телефон: 0956768062
Эл. почта: kustokust@gmail.com

О.Н. Куктенко
Телефон: 0509172327

Статья поступила 09.12.2015

Рецензент к.т.н., доц. С.В. Никульшин