
КОНСУЛЬТАЦИИ

ФАКТОРНЫЙ АНАЛИЗ: ТРАДИЦИИ ИСПОЛЬЗОВАНИЯ И НОВЫЕ ВОЗМОЖНОСТИ

Г.П.Бессокирная

(Москва)

В статье дается краткий обзор публикаций 80-90-х годов по популярному математическому методу анализа данных. На примере факторного анализа (ФА) оценок удовлетворенности различными сторонами повседневной жизни обсуждаются результаты применения традиционных и пока еще редко используемых методов выделения и вращения факторов в SPSS.

Ключевые слова: разведочный (эксплораторный) ФА, метод главных компонент, метод максимального правдоподобия, варимакс, прямой облимин.

Традиции использования

Российским социологам хорошо известны “трудные проблемы” факторного анализа (ФА) [1, с. 10-13, 40]. Однако ФА был и остается весьма популярным математическим методом анализа данных. Например, анализ публикаций журнала “Социологические исследования” с начала издания (1974 год) и до 1981 года, осуществленный Кузнецовым А.Г. и Ростегаевой Н.И., показал, что ФА применялся чаще других математических методов анализа данных [2, с. 32]. В течение 80-х годов во всех сборниках и коллективных монографиях по анализу данных, подготовленных в Отделе методического и информационного обеспечения ИСИ (ИС)

АН СССР, публиковались статьи по ФА. В работе Григорьевой Н.В. была проанализирована специфика применения ФА категориальных признаков для агрегирования социологической информации [2, с. 41-51]. Шрайбер Е.А. сравнивал факторизацию признаков с помощью многомерного шкалирования и ФА [2, с. 149-162]. Сравнительные возможности факторного и латентно-структурного анализа данных были рассмотрены Бадалян Л.Г. [3, с. 103-112]. Аргунова К.Д. и Татарова Г.Г. изучали возможности ФА при выборе стратегии анализа взаимосвязи признаков [4]. Повышенный интерес наблюдался к ФА качественных признаков [5] и особенно к проблемам интерпретации результатов ФА: статьи Дегтярева Г.П. [3, с. 168-184], Бадалян Л.Г. [6, с. 67-94], Викторова В.И. [6, с. 215-251.], Липовецкого С.С. [7].

В 90-е годы ФА по-прежнему активно используется социологами. Ядов В.А. использовал ФА для изучения социальной идентификации [8] и структуры социально-тревожного сознания [9]. В первом исследовании были выделены факторы: “вовлеченность в повседневные социальные связи”, “конформно-достигательная пассивность”, “достигательная активность”, “символические общности” [8, с. 210-214] и вычислены средние значения факторов для различных социальных групп [8, с. 215], а во втором - выделены и проанализированы факторы: “общая тревожность”, “ксенофобия- экологическая тревожность”, “катастрофизм”, “терроризм -экологическая катастрофа”, “тоталитаризм-русская самобытность”, “максимизация катастрофического сознания” [9, с. 81-84]. Магун В.С. использовал ФА для обнаружения характерных для людей естественных сочетаний различных трудовых ценностей: ФА были подвергнуты 15 суждений о субъективной важности различных аспектов работы [10]. Среди других публикаций необходимо отметить работы М.Ф.Черныша [11, 12] и В.Е.Гимпельсона [13]. Черныш М.Ф. выделяет с помощью ФА типы самоидентификации и общих политических установок, а Гимпельсон В.Е. - типы экономического сознания в

сфере занятости. Кроме того, Гимпельсон В.Е. анализирует статистическую взаимосвязь компонент экономического сознания с показателями результативности трудоустройства, вычисляет значения фактора личной ответственности и вводит значение локуса контроля в качестве переменной в регрессионную модель изменений в статусе занятости [13, с. 133-137]. В 90-е годы В.А. Ядов использовал ФА и для сравнительного анализа данных мониторинговых исследований социальной идентификации [8]. Наш опыт применения ФА связан с сопоставлением собственных результатов с результатами другого исследования по аналогичной проблематике. Сравнивались результаты ФА оценок удовлетворенности различными элементами производственной ситуации рабочих Москвы в 90-е годы и рабочих Ленинграда в 1976 году [14]. Из всего вышесказанного можно сделать вывод, что ФА, действительно, прочно вошел в арсенал математических методов анализа данных в социологических исследованиях.

Более четверти века начинающие социологи знакомились с ФА, как правило, по учебнику В.А. Ядова. Сегодня основные принципы “стандартной”, наиболее распространенной схемы применения ФА можно узнать как из учебника В.А. Ядова [15], так и из учебника по прикладной социологии [16]. История рождения ФА подробно описана И.Ф. Девятко [17, с. 122-124]. Весьма кратко принципы ФА в русле идей тестового подхода и формальный аппарат изложены в курсе лекций по измерению в социологии [18]. ФА как методу анализа данных посвящены отдельные статьи во всех социологических справочниках [19, 20]), а также главы в общеизвестных справочных пособиях по статистике (см., например [21, 22]). Особенности ФА как статистического метода анализа данных на ЭВМ освещаются в специальной литературе (см., например [23, 24, 25]). ФА представлен во всех известных статистических пакетах для ЭВМ. В распоряжении социологов имеются многочисленные программные средства, но социологи подчас не знают, как их применять.

Авторы всех вышеперечисленных публикаций, к сожалению, не объясняют, почему они используют именно те или иные методы выделения и вращения факторов. Например, Ядов В.А. в первой из двух упомянутых выше работ анализирует факторы, выделенные методом главных компонент, “после вращения”, а во второй - “до вращения”. Во второй работе, правда, отмечается, что вращение факторов методом varimax (варимакс) проясняет описанную структуру, не меняя ее природы [9, с. 85]. Гимпельсон В.Е. также использовал результаты метода главных компонент до вращения. Черныш М.Ф. в первой своей работе отмечает, что была выполнена процедура вращения матрицы по методу direct oblimin (прямой облимин). Заметим, что результаты ФА одного и того же массива данных (случайная стратифицированная выборка населения России, ноябрь 1993 года) у Черныша М.Ф. и Ядова В.А. различаются даже по числу выделенных факторов (соответственно 3 и 4). Во второй работе Черныш М.Ф. также уже не вращает факторы.

Подводя некоторые итоги анализа традиций использования ФА, можно, по-видимому, говорить о сознательном отказе социологов от вращения факторов, как одной из особенностей применения ФА в 90-е годы. Другой особенностью, с нашей точки зрения, является попытка учесть при интерпретации результатов ФА “факторную сложность переменной”, то есть ее вхождение с большими факторными нагрузками одновременно в два и более факторов [9, с. 83 и 12, с. 17]. Однако новые возможности ФА до сих пор остаются практически неиспользованными.

Новые возможности ФА

В последних версиях SPSS, к примеру, кроме традиционного метода главных компонент для выделения факторов предлагается еще шесть методов (невзвешенный МНК, обобщенный МНК, максимальное правдоподобие, факторизация главных осей, альфа

факторизация, анализ образов), а для их вращения - пять методов (кроме традиционного варимакс, прямой облимин, квартимакс, эквимакс, промакс). Какие из них выбрать?

Присоединяясь к критике бездумного использования эксплораторного ФА в качестве волшебной машинки для обнаружения “новых латентных сущностей” [17, с. 123], согласимся и с предложением Г.С. Батыгина “поиграть в факторный анализ” [26]. Игра в ФА, безусловно, нуждается в определенных правилах. В качестве таковых можно принять ориентиры в понимании методологических проблем ФА, сформулированных Б.Г. Миркиным в предисловии к монографии Петра Благуша [27, с. 5-13]. Кстати, в этой работе впервые на русском языке представлены модели и методы нелинейного ФА, приведен обширный список литературы по ФА за 20 лет (на 11 страницах). Особое внимание следует обратить на замечание Б.Г. Миркина о том, что к многочисленным недоразумениям, мифам и догмам приводит активное сосуществование в ФА алгоритмической и теоретико-вероятностной парадигм современной статистики. Метод главных компонент относится к первой, а метод максимального правдоподобия - ко второй. Нельзя не прислушаться и к мнению П.Ф. Андруковича, по работе которого, опубликованной в конце 70-х годов, многие социологи до сих пор знакомятся с ФА [28]. В начале 90-х годов он предпринял подробный разбор и переосмысление основных положений ФА. В частности, весьма актуальными, по его мнению, представляются положения о необходимости использования “замкнутой базы данных”, сопоставимых шкал и, конечно же, конструирования подходящей по содержательным соображениям модели ФА [29, с. 65].

Для более осмысленного применения ФА желательно познакомиться с выпуском серии “Quantitative Applications in Social Sciences” [30] и с работами Йореско К.Г. “Факторный анализ методами наименьших квадратов и максимального правдоподобия” [23, с. 136-168], Хармана “Метод минимальных остатков в ФА”

[23, с. 169-183], Кэттела Р.Б. и Ханна Д.К. “Принципы и процедуры однозначного поворота в ФА” [23, с. 184-218], переведенными на русский язык уже в середине 80-х годов. Следует согласиться с Ю.П.Адлером, который в предисловии к русскому переводу книги Мостеллера Ф. и Тьюки Дж., названном “Наука и искусство анализа данных”, пишет: “Единственный способ вырваться из “плена” какого-нибудь метода -это овладеть им” [31].

В настоящей статье представлены результаты ФА оценок удовлетворенности различными сторонами повседневной жизни рабочих Москвы (полевые работы были проведены в 1999 году). Из массива эмпирических данных выбраны только наблюдения без пропущенных значений переменных (329 из 546 случаев).

Нами использованы два метода выделения факторов (традиционный метод главных компонент и метод максимального правдоподобия) и два метода вращения факторов (традиционный варимакс - один из методов ортогонального вращения и прямой облимин - один из методов косоугольного вращения).

Результаты ФА

Результаты факторизации оценок удовлетворенности различными сторонами повседневной жизни обоими методами представлены в таблице 1. В ней не указаны факторные нагрузки, величина которых менее 0,300.

Анализ главных компонент (наиболее раннего метода выделения факторов) позволяет предположить, что наряду с компонентой общей удовлетворенности жизнью (в нее вошли все оценки удовлетворенности, кроме полезности полученного образования) выделяются еще две компоненты. При этом в компоненте 2 наибольшие нагрузки у двух переменных трудовых взаимоотношений, а в компоненте 3 - у оценок полезности полученного образования и величины свободного времени. Обе эти компоненты трудно интерпретировать.

ФАКТОРНЫЕ НАГРУЗКИ ДО ВРАЩЕНИЯ

Таблица 1.

Оценки	Компонента			Фактор		
	1	2	3	1	2	3
удовлетворенности:						
жизнь в целом	0,725				0,675	
структура свободного времени	0,681	-0,301			0,665	
работа на предприятии в целом	0,671				0,536	
материальное положение	0,657		0,304		0,632	
условия быта в целом	0,589		0,300		0,516	
размер зарплаты	0,575				0,456	
состояние здоровья	0,530	-0,321	-0,307		0,487	
взаимоотношения с руководством	0,319	0,744		0,999		
взаимоотношения с коллегами по работе	0,442	0,616		0,408		
полезность полученного образования			0,701			
величина свободного времени	0,436		-0,467		0,378	0,536

При использовании метода максимального правдоподобия (одного из наиболее популярных методов в ФА) были выделены: фактор 1 - удовлетворенность трудовыми взаимоотношениями, фактор 3 - удовлетворенность свободным временем, фактор 2 - общая удовлетворенность жизнью (все остальные оценки, кроме оценки полезности полученного образования). Эта оценка - кандидат на исключение в дальнейшем анализе.

Ортогональное вращение матрицы компонент (метод “варимакс”) привело к выделению (наряду с компонентой общей удовлетворенности жизнью) компоненты удовлетворенности работой (компонента 2). Вращение матрицы факторов (с помощью метода “варимакс”) также выявило по сути аналогичный этой компоненте фактор 2 (таблица 2).

РЕЗУЛЬТАТЫ ВРАЩЕНИЯ ПО МЕТОДУ “ВАРИМАКС”

Таблица 2.

Оценки	Компонента			Фактор		
	1	2	3	1	2	3
удовлетворенности:						
жизнь в целом	0,706			0,727		
структура свободного времени	0,618		0,425	0,568		0,439
работа на предприятии в целом	0,501	0,454		0,531		
материальное положение	0,733			0,678		
условия быта в целом	0,701			0,490		
размер зарплаты	0,538			0,460		
состояние здоровья	0,393		0,569	0,433		
взаимоотношения с руководством		0,818			0,999	
взаимоотношения с коллегами по работе		0,781			0,405	
полезность полученного образования	0,308		-0,658			
величина свободного времени			0,643			0,618

Особый интерес, пожалуй, представляют результаты косоугольного вращения (таблица 3).

После косоугольного вращения две первые компоненты/фактора легко интерпретируются: это общая удовлетворенность жизнью и удовлетворенность работой на предприятии. В обеих структурных матрицах две переменные (удовлетворенность жизнью в целом и удовлетворенность работой на предприятии) входят в обе компоненты/оба фактора одновременно. Кроме того, в обеих структурных матрицах компонента 3/фактор3 имеются две переменные оценки величины и структуры свободного времени, которые “склеены” с общей удовлетворенностью жизнью. Однако корреляционные матрицы компонент/факторов свидетельствуют о том, что связь трех компонент/факторов достаточно слабая (таблица 4 и таблица 5).

РЕЗУЛЬТАТЫ ВРАЩЕНИЯ ПО МЕТОДУ “ОБЛИМИН”

Таблица 3.

Оценки	Компонента			Фактор		
	1	2	3	1	2	3
удовлетворенности:						
жизнь в целом	0,730	0,358		0,302	0,726	
структура свободного времени	0,662		-0,466		0,626	0,428
работа на предприятии в целом	0,566	0,514		0,307	0,570	
материальное положение	0,731				0,658	
условия быта в целом	0,689				0,520	
размер зарплаты	0,565				0,483	
состояние здоровья	0,455		-0,593		0,457	
взаимоотношения с руководством		0,814		0,987		
взаимоотношения с коллегами по работе		0,792		0,444	0,328	
полезность полученного образования			0,636			
величина свободного времени	0,302		-0,658		0,331	0,599

КОРРЕЛЯЦИОННАЯ МАТРИЦА КОМПОНЕНТ¹

Таблица 4.

Компонента	1	2	3
1	1,00	0,22	-0,19
2	0,22	1,00	
3	-0,19		1,00

КОРРЕЛЯЦИОННАЯ МАТРИЦА ФАКТОРОВ

Таблица 5.

Компонента	1	2	3
1	1,00	0,32	-0,12
2	0,32	1,00	0,13
3	-0,12	0,13	1,00

В целом комплексное использование различных методов выделения и вращения факторов позволило получить результаты, которые с содержательной точки зрения не противоречат имеющимся

Примечание. Значения коэффициента корреляции ниже 0,100 опущены.

научным знаниям и дают возможность выбора той модели ФА, которая более адекватно отражает эмпирические данные конкретного исследования .

В нашем исследовании разные методы выделения и вращения факторов дали весьма схожие решения. Все они подтверждают существование, наряду с общим фактором удовлетворенности жизнью, относительно самостоятельного фактора удовлетворенности работой на предприятии. Это, кстати, объясняет рассогласование в динамике двух важнейших показателей социального самочувствия, зафиксированное в России в 90-е годы [14, с. 190].

Учитывая результаты предыдущего опыта анализа оценок удовлетворенности различными сторонами повседневной жизни [40], наиболее подходящей по содержательным соображениям представляется модель, которая была получена после косоугольного вращения факторов, выделенных методом максимального правдоподобия. В дальнейшем предполагается вычислить средние значения факторов по этой модели для всех респондентов и использовать их в качестве переменных при анализе социальной адаптации рабочих в трансформирующейся России.

ЛИТЕРАТУРА

1. *Гибсон У.* Факторный, латентно-структурный и латентно-профильный анализ//Математические методы в социальных науках. М.:Прогресс, 1973. С. 9-41.
2. Применение математических методов и ЭВМ в социологических исследованиях. М.:ИСИ АН СССР, 1982.
3. Комплексный подход к анализу данных в социологии. М.:ИС АН СССР, 1989.
4. *Аргунова К.Д., Татарова Г.Г.* Выбор стратегии анализа взаимосвязи признаков//Математические методы анализа и интерпретация социологических данных. М.:Наука, 1989. С. 61-94.
5. *Шашнов С.А.* О применении метода инструментальных переменных в факторном анализе качественных признаков//Математические методы в социологических исследованиях. М.: ИСИ АН СССР, 1984. С. 152-163.

6. Интерпретация и анализ данных в социологических исследованиях. М.:Наука, 1987.
7. *Липовецкий С.С.* Вопросы проведения интерпретации главных компонент//Математические методы и модели в социологии. Вып. 2. М.:ИС АН СССР, 1991. С. 43-57.
8. *Ядов В.А.* Социальные и социально-психологические механизмы формирования социальной идентичности личности//Мир России, 1995. Том IV. №3-4. С. 158-181. Мир России, 1996. Том V. №1. С. 208-216.
9. *Ядов В.А.* Структура и побудительные импульсы социально-тревожного сознания//Социологический журнал, 1997. №3. С. 77-89.
10. *Магун В.С.* Российские трудовые ценности в сравнительной перспективе//Социологические чтения. Вып 2. М.:ИС РАН, 1997. С. 134-176.
11. *Черныш М.Ф.* Национальная идентичность: особенности эволюции// Социологический журнал, 1995. №2. С. 110-114.
12. *Черныш М.Ф.* Социальная мобильность и политические ориентации// Социологический журнал, 1997. №3. С. 5-19.
13. *Гимпельсон В.Е.* Уволенные работники на рынке труда: влияние локуса контроля на восстановление занятости//Социологический журнал, 1995. №2. С. 123-140.
14. *Бессокирная Г.П., Темницкий А.Л.* Удовлетворенность работой на предприятии и удовлетворенность жизнью//Социологический журнал, 1999. №1/2. С. 186-191.
15. *Ядов В.А.* Стратегия социологического исследования. Описание, объяснение, понимание социальной реальности. М.:Добросвет, 1998. С. 345-349.
16. Основы прикладной социологии/ Под ред. Ф.Э. Шереги и М.К. Горшкова. М.:Интерпракс, 1996. С. 104-107.
17. *Девятко И.Ф.* Диагностическая процедура в социологии. М.:Наука, 1993.
18. *Толстова Ю.Н.* Измерение в социологии: Курс лекций. М.:ИНФРА-М, 1998. С. 92-104.
19. *Финкель Л.С.* Анализ факторный//Социологический справочник/ Под общ. редакцией В.И. Воловича. К.:Политиздат Украины, 1990. С. 190-193.
20. Факторный анализ/Большой толковый социологический словарь (Collins). Том 2 (П-Я). М.:Вече, АСТ, 1999. С. 380.
21. *Кендалл М., Стьюарт А.* Многомерный статистический анализ и временные ряды. М.:Наука, 1976. С. 400-436.

22. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности. М.:Финансы и статистика, 1989.
23. Статистические методы для ЭВМ/ Под ред.К.Эйслеина и др. М.:Наука, 1986.
24. SPSS Base 7.5 для Windows: Руководство по применению. М.:АО СТАСИС, 1997. С. 313-352.
25. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере/ Под ред. В.Э.Фигурнова. М.:ИНФРА-М, 1998. С. 467.
26. Батыгин Г.С., Девятко И.Ф. и др. Формирование пространства признаков в диагностической процедуре (научный доклад). М.:ИС АН СССР, 1991. С. 17.
27. Благуш П. Факторный анализ с обобщениями. М.:Финансы и статистика, 1989.
28. Андрукович П.Ф. Факторный анализ//Статистические методы анализа информации в социологических исследованиях. М.:Наука, 1979. С. 212-248.
29. Андрукович П.Ф. Заметки о факторном анализе//Многомерный статистический анализ и вероятностное моделирование реальных процессов. М.:Наука, 1990. С. 51-55.
30. Ким Дж.-О, Мьюллер Ч.У. Факторный анализ: статистические методы и практические вопросы//Факторный, дискриминантный и кластерный анализ. М.:Финансы и статистика, 1989. С. 5-87.
31. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессия. Вып. 2. М.:Финансы и статистика, 1982. С. 7.
32. Бессокирная Г.П., Темницкий А.Л. Рабочие на частном предприятии: удовлетворенность жизнью//Социологические исследования, 2000. №6. (в печати)