ТЕОРИЯ, МЕТОДОЛОГИЯ И МЕТОДЫ

DOI: 10.14515/monitoring.2018.3.02

Правильная ссылка на статью:
Зангиева И. К., Ротмистров А. Н. Сравнительный анализ способов проведения факторного анализа на порядковых переменных // Мониторинг общественного мнения : Экономические и социальные перемены. 2018. № 3. С. 29—46. https://doi.org/10.14515/monitoring.2018.3.02.

For citation:
Zangieva I. K., Rotmistrov A. N. (2018) Factor analysis of ordinal variables: a comparative study. *Monitoring of Public Opinion : Economic and Social Changes.* No. 3. P. 29—46. https://doi.org/10.14515/monitoring.2018.3.02.

И. К. Зангиева, А. Н. Ротмистров

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СПОСОБОВ ПРОВЕДЕНИЯ ФАКТОРНОГО АНАЛИЗА НА ПОРЯДКОВЫХ ПЕРЕМЕННЫХ

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СПОСОБОВ ПРОВЕДЕНИЯ ФАКТОРНОГО АНАЛИЗА НА ПОРЯДКОВЫХ ПЕРЕМЕННЫХ

ЗАНГИЕВА Ирина Казбековна — кандидат социологических наук, старший преподаватель кафедры методов сбора и анализа социологической информации, департамент социологии, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия.
E-MAIL: izangieva@hse.ru
ORCID: 0000-0001-5302-8101

FACTOR ANALYSIS OF ORDINAL VARIABLES: A COMPARATIVE STUDY

Irina K. ZANGIEVA[1] — Cand. Sci. (Soc), Assistant Professor
E-MAIL: izangieva@hse.ru
ORCID: 0000-0001-5302-8101

[1] National Research University Higher School of Economics, Moscow, Russia

РОТМИСТРОВ Алексей Николаевич — кандидат социологических наук, доцент кафедры методов сбора и анализа социологической информации, департамент социологии, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия.
E-MAIL: arotmistrov@hse.ru
ORCID: 0000-0003-2386-8710

Alexey N. ROTMISTROV[1] — Cand. Sci. (Soc), Associate Professor
E-MAIL: arotmistrov@hse.ru
ORCID: 0000-0003-2386-8710

[1] National Research University Higher School of Economics, Moscow, Russia

Аннотация. Данная работа посвящена сравнению различных способов проведения факторного анализа на порядковых данных. В исследованиях часто необходимо выявить некоторую латентную переменную, стоящую за наблюдаемыми индикаторами, измеренными по порядковой шкале. К таким индикаторам нельзя применить факторный анализ в его классическом виде, поскольку в его основе лежит матрица корреляций Пирсона, которая имеет смысл только для интервальных переменных. В этой ситуации перед исследователем встает выбор: работать с порядковыми индикаторами как с интервальными, дихотомизировать порядковые переменные или применить специальные техники, предназначенные для порядковых индикаторов: подмену матрицы корреляций или категориальный метод главных компонент (CatPCA). Посредством теоретического сравнения допущений, которые лежат в основе алгоритмов каждого из способов, а также статистического эксперимента данное исследование отвечает на вопрос, какой из перечисленных способов факторизации оптимален для выявления латентных переменных на порядковых индикаторах разной размерности — 3-, 5- и 10-балльной.

Abstract. The paper considers different approaches to the factor analysis (FA) for ordinal data. In some studies it is necessary to find a latent variable behind the observed indicators measured on an ordinal scale. Classical factor analysis cannot be applied to those indicators as it is built on the Pearson correlation coefficient which is only applicable to interval variables. So the researcher faces a choice: to treat the ordinal variables as the interval ones, to dichotomize ordinal variables or to use special techniques for ordinal indicators such as replacing the correlation matrix or using Categorical principal components analysis (CatPCA). The study is based on a theoretical comparison of assumptions that underpin the algorithms of each applications and a statistical experiment and provides an answer to the question which of the above-mentioned factorization approaches is optimal for indentifying latent variables measured by ordinal indicators on a 3-point, 5-point or 10-point scale.

Ключевые слова: латентные переменные, факторный анализ, порядковые индикаторы, метод главных компонент, категориальный метод главных компонент

Keywords: latent variables, factor analysis, ordinal variables, Principal component analysis, Categorical Principal Components Analysis

Введение

Задача свести *n* наборов наблюдаемых однородных переменных (далее — индикаторов) в n интегральных переменных, каждая из которых соответствует своему набору, часто встречается в практике анализа данных в любой науке, в которой есть место большим массивам данных. Искомая интегральная переменная может ассоциироваться как с латентной переменной в психологическом смысле — тогда речь идет о тестовой традиции (см., напр. [Толстова, 2009: гл. 7]), так и с технической процедурой снижения размерности данных без существенной потери информации (в этом смысле факторный анализ в статистическом приложении SPSS отнесен к разделу «Снижение размерности»). Далее в соответствии с традицией мы будем употреблять термин «латентная переменная» для обозначения обоих смыслов.

Самым распространенным методом факторизации является факторный анализ методом главных компонент (далее — МГК). Так, в 2005 г. исследование [Costello, Osborne, 2005: 2] показало, что только за двухлетний период примерно в 1700 опубликованных исследованиях в разных вариантах использовался факторный анализ и более чем в половине из них применялся МГК. Сегодня наблюдается похожая ситуация: согласно базе научного цитирования Scopus, в 2016 г. в области социальных наук было опубликовано 1549 статей, содержащих слова «factor analysis» в списке ключевых слов, аннотации или заглавии. Из этих 1549 статей в 46 упоминаются методы главных осей (principal axis factoring или principal axis method), максимального правдоподобия (maximum likelihood factor analysis) или наименьших квадратов (least squares); такие методы, как альфа-факторный анализ (alpha-factoring) и анализ образов (image factoring) не упоминаются вообще; в 79 статьях упоминается МГК (principal component analysis), а в остальных 1424 статьях способ получения факторной структуры не назван. Таким образом, МГК как минимум используется почти в два раза чаще, чем другие распространенные методы факторизации вместе взятые. Если же учесть, что МГК установлен в качестве метода факторизации по умолчанию во многих статистических пакетах, в том числе в SPSS и SAS [ibidem], а также что в социальных науках факторный анализ фактически отождествляется с МГК, то число статей, в которых мог использоваться МГК, достигнет 1503, что составит более 97 % от всех рассматриваемых публикаций.

МГК имеет ряд ограничений, самым важным из которых для его применения в социальных науках нам представляется то, что индикаторы должны быть метрическими (как минимум интервальными), поскольку метод предусматривает анализ матрицы коэффициентов корреляции Пирсона между индикаторами, а интерпретация этих коэффициентов имеет смысл только при метрических шкалах. В соответствии с теорией измерения, интервальной шкалу мы называем тогда и только тогда, когда любые две разности между шкальными значениями (например, a–b и c–d) можно осмысленно сравнить, то есть определить, какое утверждение содержательно правильно: a–b > c–d, a–b < c–d или a–b = c–d.

Это ограничение часто входит в противоречие с природой исходных данных в социальных науках, так как большинство социальных признаков измеряются категориальными шкалами: номинальной и порядковой, для которых процедура порождения данных не дает формальных оснований считать осмысленными соотношения расстояний между категориями шкалы. Возьмем для примера признак «удовлетворенность собственным материальным положением» с четырьмя градациями: «вполне удовлетворен», «скорее удовлетворен», «скорее не удовлетворен», «совершенно не удовлетворен». Какими бы кодами ни обозначить эти четыре градации, арифметическая разница между кодами не поддается содержательной интерпретации. Даже если градациям присвоить коды 1, 2, 3, 4, можем ли мы утверждать, что 4–3 = 2–1? Другими словами, можем ли мы утверждать, что разница степеней неудовлетворенности собственным материальным положением для людей, скорее не удовлетворенных и совершенно не удовлетворенных, равна разнице степеней удовлетворенности для людей, скорее удовлетворенных и вполне удовлетворенных? — Не можем. Как не можем утверждать, что первая разница больше, чем вторая, или что первая разница меньше, чем вторая. Сравнивать их корректно мы просто не можем, исходя из процедуры приписывания градациям степени удовлетворенности числовых индексов.

Для контраста приведем пример признака, также характеризующего материальное положение, но для которого есть формальное основание считать осмысленными соотношения расстояний между категориями шкалы; речь о месячном доходе в денежном выражении. Очевидно, что формально разница между доходами 5000 руб. и 10000 руб. равна разнице между доходами 105000 руб. и 110000 руб. (Хотя если посмотреть содержательно и рассмотреть месячный доход как «признак-прибор» [Толстова, 2000: ч. 2, п. 1.1, ч. 1, п. 1.3], все будет не так очевидно.)

Настоящая статья посвящена поиску оптимальных путей примирения противоречия между шкальными требованиями МГК и природой данных социальных исследований в различных исследовательских ситуациях, под которыми здесь мы понимаем различную размерность (число градаций) примененных порядковых шкал.

Способы применения МГК к порядковым индикаторам: суть и априорное сравнение

В практике и теории анализа данных получили распространение несколько способов применения МГК к порядковым индикаторам:

1. принять (волевым образом) допущение о том, что порядковые индикаторы можно считать интервальными [Knapp, 1990];
2. дихотомизировать порядковые индикаторы, считая дихотомические переменные частным случаем интервальных [Vermunt, Magidson, 2005];
3. вместо матрицы коэффициентов корреляции Пирсона использовать матрицу коэффициентов ранговой или полихорической корреляции [Rigdon, Ferguson, 1991];

4. применить категориальный метод главных компонент (далее — CatPCA), в рамках которого МГК применяется к специально преобразованным индикаторам [Meulman, Kooij, Heiser, 2004].

Выбор способа обусловлен характером исследовательской ситуации.

Чтобы сузить рамки экспериментального исследования, мы провели теоретическое сравнение приведенных четырех вариантов факторизации порядковых индикаторов посредством МГК по следующим критериям:

— требования к данным (тип шкалы, важность формы распределения),
— характер связи — как между самими индикаторами, так и между индикаторами и фактором (линейный или монотонный),
— необходимость преобразования исходных индикаторов.

Допуская, что порядковые индикаторы, измеренные по шкале определенной размерности, можно считать интервальными, исследователь тем самым соглашается и с тем, что они в целом удовлетворяют всем **требованиям МГК**: что порядковые индикаторы непрерывны, а распределения их признаков симметричны относительно среднего арифметического.

Выбирая дихотомизацию, исследователь ориентируется на аналогичные требования, но применительно к дихотомическим переменным. Специфика состоит в том, что среднее арифметическое дихотомической переменной, закодированной «0» и «1», интерпретируется как доля объектов, имеющих значение «1» среди всех изученных объектов [Толстова, 2000: ч. 2, п. 1.2]. При такой трактовке среднего арифметического предъявляемое МГК требование симметричности распределения превращается в требование его равномерности. А непрерывность дихотомических переменных проявляется в способности характеризовать совокупность изучаемых объектов любой точкой, отражающей (якобы) среднее арифметическое на отрезке [0;1].

При отказе от исходной матрицы коэффициентов корреляции Пирсона в пользу матрицы ранговых или полихорических коэффициентов, или в пользу способа CatPCA мы облегчаем требования к подлежащим факторизации индикаторам: достаточно, чтобы все индикаторы измерялись шкалой, не ниже порядковой, а для CatPCA такой индикатор должен быть один.

Таким образом, выбирая способ 1 или 2, исследователь вынужден принимать серьезные допущения и быть гораздо более осмотрителен, чем при выборе способа 3 или 4. Такая осмотрительность не всегда проявляется, что находит критический отклик в академической литературе. В частности, критикуют принятие порядковой шкалы за интервальную без необходимого учета характеристик распределения переменной и величины выборки [Knapp, 1990], это даже назвали одним их «сети грехов статистического анализа» [Kuzon et al., 1996: 265].

За исключением использования матрицы ранговых или полихорических коэффициентов, все способы факторизации предполагают линейный **характер связи** как между самими индикаторами, так и между индикаторами и факторами, поскольку эта связь измеряется коэффициентом корреляции Пирсона.

Специфика CatPCA заключается в том, что для построения матрицы коэффициентов корреляции Пирсона нужно преобразовать исходные индикаторы. Поэтому, строго говоря, в рамках CatPCA характер связи между исходными индикаторами и между исходными индикаторами и факторами может быть любым, в том числе точечно-линейным (когда связаны лишь отдельные категории пар индикаторов и пар индикаторов и факторов), и роль CatPCA как раз в том и состоит, чтобы преобразовать любую нелинейную связь в линейную. Такое преобразование базируется на модельном предположении CatPCA о том, что в сознании респондентов имеются непрерывные латентные переменные, линейно выражаемые исходными индикаторами (т. е. равные отрезки этих исходных индикаторов отражают примерно равное изменение в степени выраженности соответствующей латентной переменной), но грубость сделанных измерений искажает линейный характер связи. Это модельное предположение CatPCA перебивает модельное предположение МГК об интервальности индикаторов и о линейности связи между ними только в одном аспекте: в гипотетическом постулировании грубости измерений исходных индикаторов. Однако трудно спорить с тем, что порядковые индикаторы действительно менее точно отражают измеряемый объект, чем интервальные. Поэтому модельное предположение CatPCA сработает только в том случае, если латентные переменные действительно непрерывны и переход каждого индикатора от одной градации к другой отражается в изменении степени выраженности соответствующей латентной переменной.

При использовании вместо матрицы коэффициентов корреляции Пирсона матрицы ранговых или полихорических коэффициентов корректировать наблюдаемые индикаторы под модельные предположения МГК не требуется; вместо корректировки индикаторов постулируется отказ от предположения о линейности рассматриваемых связей. Другими словами, отказ от исходной матрицы коэффициентов корреляции Пирсона делает ненужным учет соотношения расстояний между рангами индикаторов, так как вовсе не обязательно, чтобы, скажем, равные расстояния между градациями индикаторов хотя бы примерно отражали равное изменение в степени выраженности искомых латентных переменных. Поскольку для порядковых индикаторов такого рода расстояния отчасти теряют смысл, то следует либо этот смысл восстановить, сделав выбор в пользу CatPCA, либо применить такой метод факторизации, в котором точные расстояния между значениями индикаторов огрубляются (как в матрицах ранговой или полихорической корреляции). Правда, замена корреляционной матрицы Пирсона полихорической корреляционной матрицей критикуется [Flora, Curran, 2004], поскольку допущение о нормальности распределения латентных переменных, стоящих за индикаторами, на практике трудно проверить; там же описываются специальные техники замены матрицы коэффициентов корреляции Пирсона на матрицу ранговую или полихорическую корреляционную матрицу.

Преобразование исходных индикаторов происходит при движении, как по пути дихотомизации, так и по пути применения CatPCA, но очень по-разному. Дихотомизация делается вручную и сводится к сворачиванию многоместной шкалы присвоением исходным индикаторам одного из двух кодов — «0» или «1». Она сопряжена с рядом трудностей, первая из которых: где установить границу между нулевым и единичным значениями? Вернемся к примеру с признаком «удовлетворенность собственным материальным положением», но допустим, что он имеет три градации: «высокая», «средняя», «низкая». Должна ли пройти граница между нулевым и единичным значениями так, чтобы отделить высокую удовлетворенность от средней и низкой? Или лучше отделить низкую удовлетворенность от средней и высокой? С содержательной точки зрения ни один вариант не представляется оптимальным. При невозможности найти содержательно оптимальный вариант можно поискать вариант, оптимальный для удовлетворения требований МГК, а именно упомянутого выше требования равномерности распределения. То есть из вариантов «высокая удовлетворенность / средняя и низкая удовлетворенность» и «высокая и средняя удовлетворенность / низкая удовлетворенность» выбрать вариант с более равномерным распределением частот.

Вторая трудность следует из первой: можно ли применять к разным индикаторам разные принципы дихотомизации? Скажем, если в батарее 3-ранговых индикаторов с градациями «высокая, средняя, низкая», то можно ли для одних индикаторов проводить границу между «высокая» и «средняя и низкая», а в других — между «высокая и средняя» и «низкая»? Наука не дает однозначного ответа. Как видно, процедура дихотомизации сопряжена с большой свободой действий исследователя. Хорошо, если выбор удается обосновать и учесть специфику выбранного принципа дихотомизации при интерпретации результатов факторизации. Но обоснование и учет могут быть трудны и затратны по времени, поэтому такая свобода может показаться избыточной, переходящий в волюнтаризм.

Совершенно иначе производят преобразования исходных индикаторов в рамках CatPCA, который включает в себя развитый алгоритм оцифровки. Можно вывести следующую логическую формулу:

CatPCA = оцифровка исходных индикаторов + МГК.

Эта формула демонстрирует, что ключом к пониманию метода CatPCA как раз является заложенная в нем процедура оцифровки. Ее суть состоит в таком изменении расстояний между категориями исходных индикаторов, чтобы связь между оцифрованными индикаторами и факторами стала максимально линейной. Оцифровка — итеративный процесс, идущий независимо от исследователя. Конечно, при желании исследователь может вмешиваться в оцифровку, принудительно задавая число категорий в оцифрованных индикаторах и число итераций. Но такое вмешательство несет дополнительные ограничения для процесса оцифровки, что обычно снижает математическое качество получаемой факторной модели. Поэтому для вмешательства в процесс оцифровки нужны веские содержательные основания.

Любопытно, что оцифровка в CatPCA нередко приводит к дихотомизации исходных индикаторов. В таком случае дихотомизация индикаторов с последующим применением традиционного МГК и CatPCA как будто бы дают похожие результаты. Однако высокое математическое качество итоговой факторной модели при дихотомизации гарантируется только квалификацией исследователя (правильно выбрал принцип дихотомизации, правильно перекодировал), тогда как в CatPCA квалификацию исследователя (понимание исследователем сути оцифровки и правильный выбор ее для соответствующих исходных данных) дополняет гарантированное качество автоматической процедуры. Из сказанного следует, что качественно сделанная вручную дихотомизация является частным случаем применения CatPCA. В то же время ряд исследователей считает сомнительным применение МГК к дихотомическим переменным [Bartholomew et al., 2002; Suits, 1957], и тогда все эти преобразования исходных данных бесполезны.

Таблица 1. ***Способы применения МГК к порядковым индикаторам***

Способы	Преобразование исходных порядковых индикаторов	Замена коэффициента корреляции Пирсона на ранговые или полихорические коэффициенты	МГК
1. Принять допущение о том, что порядковые индикаторы, измеренные по шкале определенной балльности, можно считать интервальными.	–	–	+
2. Дихотомизировать индикаторы, склеив соседние градации в переменные со значениями 0 и 1 для дальнейшего применения МГК, считая дихотомические переменные частным случаем интервальных.	+	–	+
3. Замена исходной матрицы коэффициентов корреляции Пирсона для индикаторов на матрицу коэффициентов ранговой или полихорической корреляции	–	+	+
4. CatPCA	+	–	+

Как видно из сравнения, ведя к одной и той же цели, рассматриваемые способы оказываются довольно несхожими (см. табл. 1), и относительно большинства из них выдвигаются критические аргументы. Поэтому выбор адекватного способа факторизации применительно к конкретной ситуации не является автоматическим. Далее мы приведем обоснования практическим рекомендациям по факторизации порядковых индикаторов посредством МГК с учетом одной их ключевых характеристик исследовательской ситуации — числа градаций индикаторов. Мы выясняли, какой способ из четырех приводит к лучшему результату применения МГК к 3-, 5- и 10-балльным индикаторам.

Такая балльность шкал выбрана неслучайно. Согласно нашей гипотезе, 10-балльная порядковая шкала «ведет» себя почти как интервальная (то есть удовлетворяет перечисленным выше требованиям МГК); 3-балльные индикаторы как наименее дробные из возможных порядковых индикаторов демонстрируют «поведение», наименее похожее на «поведение» интервальной шкалы (максимально «грубые» индикаторы); 5-балльная шкала выбрана как промежуточный вариант. Эта гипотеза опирается на психофизиологические исследования (см. работы Луи Терстоуна и его последователей) и практический опыт социальных измерений. Исследователю нужно понимать, имеются ли в сознании каждого респондента расстояния между значениями изучаемых признаков и может ли он хотя бы примерно соотносить эти расстояния между собой: если имеются и может, то это как минимум интервальный тип шкалы; если нет — то категориальный. Известна следующая предпосылка соотносимости рассматриваемых расстояний в сознании респондентов: шкала вопроса имеет большое число (от 7) упорядоченных градаций, выражающих степень интенсивности некоторого свойства, скажем, удовлетворенность работой (обоснование см., напр., [Толстова, 2009: пп. 1.1, 10.1—11.2]), и/или имеет единицы измерения (часов в неделю, раз в день, рублей, штук и т. п.). При наличии этой предпосылки шкалу можно считать интервальной. Если же значений у шкалы немного (менее семи) и/или они выражают не степень интенсивности некоторого свойства, а самостоятельные состояния (скажем, тип трудового договора), то это порядковая шкала.

Специальные исследования показали, что коэффициент корреляции Пирсона плохо работает для шкал с менее чем пятью градациями [Bollen, Barb, 1981; O'Brien, 1979]. Как следствие сходства 10-балльной шкалы с интервальной шкалой, мы предполагаем, что для 10-балльных индикаторов будет оптимальным считать их интервальными, тогда как для 3- и 5-балльных индикаторов оптимальным окажутся CatPCA или подмена исходной матрицы коэффициентов корреляции Пирсона на матрицу коэффициентов ранговой полихорической корреляции. Нашу гипотезу мы проверили посредством статистического эксперимента.

Дизайн эксперимента

Для сравнительной оценки работоспособности предлагаемых способов факторизации порядковых индикаторов посредством МГК был проведен статистический эксперимент. Его суть заключалась в том, чтобы сгенерировать (эталонный) массив из сугубо интервальных нормально распределенных переменных; получить на них факторную модель; преобразовать эталонный массив в три экспериментальных массива путем преобразования интервальных переменных в переменные с 3, 5 и 10 градациями соответственно; на каждом экспериментальном массиве получить факторные модели с помощью четырех способов факторизации; сравнить полученные на экспериментальных массивах модели с эталонной моделью. Наиболее подходящим способом факторизации порядковых индикаторов признается тот, результаты которого в наибольшей степени будут похожи на эталонные результаты.

Генерирование эталонного массива и построение эталонной модели

Чтобы быть уверенными, что результаты эксперимента характеризуют ситуации, различающиеся именно размерностью индикаторов, а не содержательной спецификой данных, мы работали не на реальном массиве данных, а на специально сгенерированных: эталонном и экспериментальных массивах.

Эталонный массив содержал 12 интервальных переменных с такой структурой корреляций, которая предполагает наличие латентных переменных; а именно содержит пучки переменных, так что переменные, принадлежащие к одному и тому же пучку, скоррелированы тесно, а переменные, принадлежащие к разным пучкам, почти не скоррелированы.

Генерируя эталонный массив, мы опирались на Renald's SPSS Tools [1]:

1. Выбрали число индикаторов и факторов в типичном для социологии соотношении: 12 индикаторов к четырем факторам; по три индикатора на фактор.
2. Сгенерировали 12 непрерывных, нормально распределенных переменных вероятностным образом. Оказалось, что переменные коррелировали между собой, хотя и очень слабо.
3. Чтобы избавиться от этой скоррелированности, мы предварительно применили МГК (для решения чисто технической задачи получения абсолютно нескоррелированных переменных) и получили 12 технических факторных переменных — непрерывных, нормально распределенных, стандартизированных и, что важно на этом шаге, ортогональных относительно друг друга.
4. Создали матрицу корреляций (см. табл. 2), которая соответствует запланированной структуре пучков индикаторов, и согласно ей преобразовали полученные на предыдущем шаге переменные. Для этого к ним применили разложение Холецкого (о нем подробнее см. [Press et al., 1992]) — метод, позволяющий связать вектор X из независимых стандартных нормальных случайных величин с желаемой корреляционной матрицей *A* = *L LT* (где *L* — нижняя треугольная матрица с положительными действительными элементами на диагонали, а *LT* — зеркальное относительно главной диагонали отображение матрицы *L*; она же верхняя треугольная матрица). В нашем случае имеется ряд векторов $\mathbf{X}_1$, $\mathbf{X}_2$, …, $\mathbf{X}_n$, соответствующих строкам эталонного массива (или объектам наблюдения в реальном исследовании).

Таблица 2. ***Матрица коэффициентов корреляции для сгенерированных данных***

Переменные	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1	0,9	0,3	0,2	0,1	0,1	0,1	0,8	0,1	0,1	0,1	0,1
2	0,9	1	0,4	0,3	0,2	0,1	0,1	0,8	0,1	0,1	0,1	0,1
3	0,3	0,4	1	0,9	0,8	0,1	0,1	0,1	0,1	0,1	0,1	0,1
4	0,2	0,3	0,9	1	0,9	0,1	0,1	0,1	0,1	0,1	0,1	0,1
5	0,1	0,2	0,8	0,9	1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
6	0,1	0,1	0,1	0,1	0,1	1	0,9	0,1	0,9	0,1	0,1	0,1
7	0,1	0,1	0,1	0,1	0,1	0,9	1	0,1	0,9	0,1	0,1	0,1
8	0,8	0,8	0,1	0,1	0,1	0,1	0,1	1	0,1	0,1	0,1	0,1
9	0,1	0,1	0,1	0,1	0,1	0,9	0,9	0,1	1	0,1	0,1	0,1
10	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	1	0,9	0,8
11	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,9	1	0,8
12	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,8	0,8	1

[1] Renald's SPSS Tools. (2005) Генерация многомерного нормального вектора с заданной ковариационной матрицей [Электронный ресурс]. 2005. Dec. 29. URL: http://www.spsstools.net/ru/syntax/syntax-index/bootstrap-and-random-numbers/generating-multivariate-normal-variables-with-a-specific-covariance-matrix (дата обращения 20.07.2017).

5. Применили МГК к эталонному массиву для получения эталонной модели, с которой и сравнивались все последующие. Количество факторов устанавливалось по критерию Кайзера, вращение осуществлялось методом Варимакс. Судя по мере адекватности КМО, критерию Бартлетта (см. табл. 3) и полной объясненной дисперсии (см. табл. 4), эталонная модель получилась высокого качества.

Таблица 3. ***Мера адекватности и критерий Бартлетта эталонной модели***

Мера выборочной адекватности Кайзера-Мейера-Олкина		,746
Критерий сферичности Бартлетта	Прибл. хи-квадрат	13073,631
	Ст.св.	66
	Знч.	,000

Таблица 4. ***Полная объясненная дисперсия эталонной модели***

Компонента	**Начальные собственные значения**			**Суммы квадратов нагрузок извлечения**			**Суммы квадратов нагрузок вращения**		
	Итого	**% Дисперсии**	**Кумулятивный %**	**Итого**	**% Дисперсии**	**Кумулятивный %**	**Итого**	**% Дисперсии**	**Кумулятивный %**
1	3,793	31,608	31,608	3,793	31,608	31,608	2,8	23,332	23,332
2	2,581	21,511	53,118	2,581	21,511	53,118	2,76	22,997	46,329
3	2,415	20,121	73,239	2,415	20,121	73,239	2,676	22,299	68,628
4	2,116	17,629	90,869	2,116	17,629	90,869	2,669	22,241	90,869
5	0,317	2,646	93,514						
…									
11	0,082	0,685	99,459						
12	0,065	0,541	100						

Как видно из таблицы 5, интервальные индикаторы распределились поровну по четырем факторам, как и планировалось при генерации эталонного массива.

Таблица 5. ***Матрица факторных нагрузок после вращения эталонной модели***

	Компоненты			
	1	**2**	**3**	**4**
nr1	,047	,106	,949	,049
nr2	,043	,217	,936	,044
nr3	,045	,929	,182	,046
nr4	,048	,971	,099	,049
nr5	,054	,939	,025	,055
nr6	,962	,048	,049	,051
nr7	,962	,048	,049	,051
nr8	,056	-,007	,918	,058
nr9	,962	,048	,049	,051
nr10	,049	,048	,048	,952
nr11	,049	,048	,048	,952
nr12	,051	,050	,050	,913

Преобразование сгенерированных интервальных переменных в порядковые индикаторы разной размерности; построение выборочных моделей

Второй этап эксперимента заключался в создании экспериментальных массивов, состоящих из порядковых переменных. Три экспериментальных массива были получены из эталонного посредством перекодирования 12 исходных интервальных переменных в порядковые — назовем эти экспериментальные массивы **корневыми**.

При создании корневого 3-балльного массива 12 исходных интервальных переменных были разбиты на три равных интервала; при создании 5-балльного — на 5 равных интервалов; при создании 10-балльного — на 10 равных интервалов. Таким образом мы получили три корневых экспериментальных массива, каждый из которых содержит по 12 трех-, пяти- и десятибалльных переменных.

Затем из каждого корневого массива мы сгенерировали три семейства **выборочных** массивов (по 15 в каждом) на основе вероятностного принципа; к каждому выборочному массиву были применены все четыре изучаемых способа факторизации. Были получены 12 семейств выборочных факторных моделей (4 изучаемых способа факторизации умножить на 3 семейства выборочных массивов). Наличие семейства факторных моделей, относящихся к одному и тому же способу факторизации индикаторов из одного и того же экспериментального массива позволило оценить устойчивость характеристик этих моделей. Причем для дихотомизации порядковых индикаторов было создано от 30 до 45 выборочных массивов, учитывая трудности, о которых говорилось выше. А именно дихотомизация 3-балльных индикаторов была осуществлена в двух вариантах: в одном объединение градаций «1» и «2» противопоставлялось градации «3»; во втором объединение градаций «3» и «2» противопоставлялось градации «1». Дихотомизация 5-балльных индикаторов была также осуществлена в двух вариантах: в одном объединение градаций «1», «2» и «3» противопоставлялось объединению градаций «4» и «5», во втором объединение градаций «3», «4» и «5» противопоставлялось объединению градаций «1» и «2». Дихотомизация 10-балльных индикаторов была осуществлена в трех вариантах: (1–5) vs. (6–10), (1–3) vs. (4–10) и (1–7) vs. (8–10).

В таблице 5 представлены значения характеристик факторных моделей (мера адекватности КМО, критерий Бартлетта и полная объясненная дисперсия) для каждого способа факторизации, усредненные в рамках каждого семейства выборочных массивов.

Таблица 6. ***Усредненные характеристики факторных моделей, полученных на экспериментальных массивах***

3-балльная шкала		**Эталонный массив**	**Принятие интервальности**	**Дихотомизация**		**Подмена матрицы корреляций**	**CatPCA**
				1+2, 3	**1, 2+3**		
Мера выборочной адекватности КМО		0,746	0,712	0,684	0,692	0,709	0,73
Критерий сферичности Бартлетта	Прибл. хи-квадрат	13073	5017	4040	4043	4595	5166
	Знч.	0,000	0,000	0,000	0,000	0,000	0,000
% объясненной дисперсии		90,9	73,3	72,8	72,4	75,4	80,1
Коэффициент конгруэнтности		—	0,996	0,997	0,997	0,98	1
КФА	RMSEA	—	0,039	**0,028**	**0,240**	0,038	0,300
	CFI	—	0,985	**0,991**	**0,993**	0,985	0,990
5-балльная шкала		**Эталонный массив**	**Принятие интервальности**	**Дихотомизация**		**Подмена матрицы корреляций**	**CatPCA**
				1–2, 3–5	**1–3, 4–5**		
Мера выборочной адекватности КМО		0,746	0,741	0,701	0,692	0,709	0,741
Критерий сферичности Бартлетта	Прибл. хи-квадрат	13073	7177	4988	4681	4595	7035
	Знч.	0,000	0,000	0,000	0,000	0,000	0,000
% объясненной дисперсии		90,9	81,6	76,1	76,2	81,6	82,3
Коэффициент конгруэнтности		—	1	0,998	0,998	1	0,999
КФА	RMSEA	—	0,060	**0,0190**	**0,020**	0,0470	0,0470
	CFI	—	0,975	**0,997**	**0,997**	0,984	0,985

10-балльная шкала		**Эталонный массив**	**Принятие интервальности**	**Дихотомизация**			**Подмена матрицы корреляций**	**CatPCA**
				1–5, 6–10	**1–3, 4–10**	**1–7, 8–10**		
Мера выборочной адекватности КМО		0,746	0,746	0,727	0,692	0,681	0,746	0,754
Критерий сферичности Бартлетта	Прибл. хи-квадрат	13073	10428	5942	4060	3923	9983	10425
	Знч.	0,000	0,000	0,000	0,000	0,000	0,000	0,000
% объясненной дисперсии		90,9	87,8	77,4	71,6	69,7	87,2	90,2
Коэффициент конгруэнтности		—	1	0,999	0,995	0,995	1	0,995
КФА	RMSEA	—	0,085	**0,038**	**0,021**	**0,022**	0,075	0,080
	CFI	—	0,966	**0,988**	**0,995**	**0,994**	0,973	0,971

Сравнение полученных экспериментальных моделей с эталонной

12 семейств выборочных экспериментальных факторных моделей были получены эксплораторным путем (поскольку МГК обычно применяется в эксплораторной роли), но, согласно идее эксперимента, они должны выражать исходную факторную структуру, заложенную при генерировании эталонного массива, а не какую-то иную факторную структуру, которая гипотетически могла бы быть найдена в выборочных массивах.

Поэтому прежде чем оценивать качество выборочных факторных моделей, первым делом мы должны были удостовериться, что выборочные факторные массивы воспроизвели исходную факторную структуру, заложенную при генерировании эталонного массива. Причем поскольку большинство из четырех изучаемых способов факторизации предполагает дополнительное преобразование данных, то проверке подлежали все выборочные массивы. Для этого мы применили **конфирматорный факторный анализ** (далее — КФА). КФА сравнивает наблюдаемую дисперсионно-ковариационную матрицу («Наблюдаемой называется дисперсионно-ковариационная матрица ковариаций индикаторов, не ограниченная предположениями о существовании факторной модели» [Назаров, Мальцев, 2006: 77]) и ожидаемую (в нашем случае соответствующую в эталонной модели). Наиболее распространенными показателями для сравнения матриц являются производные от критерия «Хи-квадрат» показатели:

$$\text{Normed Comparative Fit Index (CFI)} = 1-(\chi^2_e-df_e)/(\chi^2_o-df_o) \qquad (1),$$

где χ^2_o — величина критерия «Хи-квадрат» для базовой матрицы (в нашем случае наблюдаемой),

df_o — число степеней свободы этого критерия;

χ^2_e — величина критерия «Хи-квадрат» для ожидаемой матрицы (в нашем случае эталонной),

df_e — число степеней свободы этого критерия; интерпретируемый диапазон изменений [0;1]; чем выше коэффициент, тем меньше различие матриц;

Root Mean Squared Error of Approximation, (RMSEA) = $\sqrt{\max\left(\chi^2{}_e - df_e, 0\right) / df_e(n-1)}$ (2),
где χ^2_e — величина критерия «Хи-квадрат» для ожидаемой матрицы (в нашем случае эталонной),

df_e — число степеней свободы этого критерия; интерпретируемый диапазон изменений [0;1]. Чем меньше коэффициент, тем меньше различие матриц; приемлемой считается величина ниже 0,6.

Как видно из таблицы 6, усредненные показатели КФА для каждого семейства выборочных массивов находятся на приемлемом уровне. Известный автор работ по мерам качества моделей КФА Д. А. Кенни считает, что показатель RMSEA первичен [2], и если он ниже 0,158, то CFI не улавливает различий в качестве моделей КФА. В нашем исследовании усредненный RMSEA только для 3-балльных массивов с дихотомизированными индикаторами и с индикаторами, подготовленными в рамках CatPCA, превысил обозначенный порог. Отсюда мы сделали вывод, что на этих массивах «труднее» воспроизвести эталонную факторную структуру, чем на прочих. Однако и для этих массивов усредненные показатели КФА находятся на приемлемом уровне, то есть и эти выборочные массивы, и, тем более, остальные воспроизводят в себе закономерности, которые были заложены на этапе генерирования эталонного массива в виде эталонной корреляционной матрицы. Поэтому можно ожидать, что 12 семейств экспериментальных массивов могут дать факторные модели, выражающие именно эталонную факторную структуру.

Чтобы подтвердить данное предположение, следовало также убедиться, что структура факторных нагрузок экспериментальных моделей действительно очень близка структуре факторных нагрузок эталонной моделей. Для этого был применен **коэффициент конгруэнтности** (далее КК), предложенный в 1948 г. Бертом и доработанный Такером [Abdi, 2007]. КК позволяет определить соответствие структуры факторов двух факторных решений по формуле:

$$\phi_{qp} = \frac{\sum_{j=1}^{n} (1^a\, jp * 2^a\, jp)}{\sqrt{\sum_{j=1}^{n} (1^a\, jp) \sum_{j=1}^{n} (2^a\, jp)}} \qquad (3),$$

где p — число факторов одной модели (в нашем случае — эталонной),

q — число факторов другой модели (в нашем случае — одной из экспериментальных),

1^a_{jp} — факторная нагрузка j индикатора с фактором p в эталонной модели,

2^a_{jp} — факторная нагрузка j индикатора с фактором q в рассматриваемой экспериментальной модели. Коэффициент изменяется в интервале [–1;1], где значение –1 говорит о полном несоответствии факторных структур, а 1 — о полном их соответствии. Применяя КК, стоит внимательно относиться к порядку факторов в полученном решении, поскольку зачастую распределение индикаторов по фак-

[2] Kenny D. A. (2015) Measuring Model Fit [Электронный ресурс]. 2015. Nov. 24. URL: http://davidakenny.net/cm/fit.htm (дата обращения 29.10.2017).

торам различается только за счет порядка самих факторов, что с содержательной точки зрения не несет в себе никаких различий.

При применении к порядковым массивам каждого из способов факторизации было получено по четыре фактора, все модели, согласно величинам КК, продемонстрировали соответствие эталонной факторной структуре (см. табл. 6).

Далее мы сравнили экспериментальные факторные модели между собой в рамках семейств массивов каждой балльности, опираясь на типичную для МГК характеристику — полную объясненную дисперсию (см. табл. 6).

3-балльные порядковые шкалы. Способ принятия интервальности 3-балльной шкалы и непосредственного применения МГК к ней дает средний процент объясненной дисперсии 73 % (почти на 20 % меньше эталонного), то есть объяснительная способность модели заметно ухудшилась, однако все еще достаточно высока (согласно практике применения МГК). Оба варианта дихотомизации дали средний процент объясненной дисперсии 72 %. Замена матрицы корреляций Пирсона на матрицу корреляций Спирмена перед применением МГК дала средний процент объясненной дисперсии 75 %. CatPCA показал наилучший результат: средний процент объясненной дисперсии составил 80,1 %, что значительно ближе к эталонной модели.

5-балльные порядковые шкалы. Здесь средний процент объясненной дисперсии факторных моделей, полученных каждым из четырех изучаемых способов факторизации, ожидаемо вырос по сравнению с факторными моделями на 3-балльных индикаторах. Лучший результат снова показал CatPCA (82,3 %), но близкий результат получен и способом, основанным на принятии интервальности 5-балльной шкалы.

10-балльные порядковые шкалы. Средний процент объясненной дисперсии факторных моделей, полученных каждым из четырех изучаемых способов факторизации, снова ожидаемо вырос по сравнению с факторными моделями на 5-балльных индикаторах. Исключение — способ дихотомизации, для которого средний процент объясненной дисперсии не вырос. Лучший результат снова показал CatPCA, причем этот результат уже почти равен эталонному.

Выводы и заключение

Мы наглядно убедились, что чем ближе по своему типу шкалы порядковые индикаторы к интервальным, тем лучше «работает» на них МГК. При этом даже «начальная точка» этого приближения (то есть массивы с 3-балльными индикаторами) дают приемлемые относительно эталонной модели результаты. Есть гипотеза, что сильное изменение распределения порядковых индикаторов в сравнении с переменными эталонного массива (прежде всего сильное отклонение распределения от симметричного относительно релевантной меры центральной тенденции) приведет к тому, что применение МГК к массивам с 3-балльными индикаторами не будет давать приемлемые результаты. В своем исследовании мы ограничились использованием схожих (симметричных) распределений порядковых индикаторов и переменных эталонного массива; чтобы сделать более общий вывод о приемлемости применения МГК к порядковым переменным «как есть», следует ввести в исследование переменные с сильно смещенными распределениями.

Кроме того мы выявили, что степень приближения результатов МГК на порядковых индикаторах к эталонной модели (чем выше число баллов порядковых индикаторов, тем ближе результаты МГК на этих переменных к эталонным результатам) разная для разных способов подготовки порядковых индикаторов для проведения МГК: наилучшую «траекторию» приближения показывает CatPCA (от 80,0% объясненной дисперсии на 3-балльных индикаторах к 90,2% объясненной дисперсии на 10-балльных индикаторах), наихудшую — способ дихотомизации (от 72,6% средней объясненной дисперсии на 3-балльных индикаторах до 72,9% на 10-балльных индикаторах). Пожалуй, начиная с 5-балльных массивов, приемлемое приближение (с учетом простоты реализации) к эталонной модели демонстрирует способ простого принятия интервальности 5-балльных индикаторов. Правда, скорее всего, это тоже обусловлено сходством распределений порядковых индикаторов и переменных эталонного массива (и те, и другие имеют один пик и симметричны относительно релевантных мер центральной тенденции).

Несмотря на то, что дихотомизация является частным случаем CatPCA, в эксперименте эти способы дали противоположные по своему качеству результаты. Видимо, сконструированные нами ситуации плохо подходят для дихотомизации и, возможно, поэтому при преобразовании исходных индикаторов CatPCA не производил дихотомизацию. Полагаем, что при каких-то иных распределениях исходных порядковых индикаторов CatPCA в результате преобразования мог бы на выходе дать дихотомию; вероятно, в таких ситуациях и дихотомизация вручную дала бы хорошие результаты.

Таким образом, результаты сравнений позволили сформулировать следующие рекомендации по использованию различных способов факторизации порядковых данных (см. табл. 7), где «/» означает равноценность способов для индикатора указанной размерности.

Таблица 7. ***Рекомендации по подготовке порядковых индикаторов для факторизации посредством МГК***

Размерность индикатора	Способ применения МГК
3-балльная	CatPCA
5-балльная	CatPCA/принятие интервальности
10-балльная	CatPCA/принятие интервальности

Список литературы (References)

Назаров Б., Мальцев В. Структурные ковариационные модели в социологии: Учебное пособие. М. : ГУ–ВШЭ. 2006.
Nazarov B., Maltsev V. (2006) Structural covariance models in sociology: Textbook. Moscow: SU-HSE. (In Russ.)

Толстова Ю. Н. Анализ социологических данных. М. : Научный мир. *2000.*
Tolstova Yu.N. (2000) Analysis of sociological data. Moscow: The scientific world. (In Russ.)

Толстова Ю. Н. Измерение в социологии : учебное пособие. 2-е изд., перераб. и доп. М. : КДУ. 2009.
Tolstova Yu.N. (2009) Measurement in Sociology: A Training Manual. 2-nd ed., Rev. and add. Moscow.: KDU. (in Russ.)

Abdi H. (2007) RV coefficient and congruence coefficient. In: *N.J. Salkind (Ed.): Encyclopedia of Measurement and Statistics*. Thousand Oaks (CA): Sage. P. 849—885.

Bartholomew D. J., Steele F., et al. (2002) The Analysis and Interpretation of Multivariate Data for Social Scientists. Bristol: CRC Press. URL: http://www.bristol.ac.uk/cmm/team/fs/aimdss-2nd-ed/chapter8.pdf (accessed 20.07.2017).

Bollen K. A., Barb K. H. (1981) Pearson's r and coarsely categorized measures. *American Sociological Review*. Vol. 46. P. 232—239. https://doi.org/10.2307/2094981.

Costello A., Osborne J. (2005) Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*. Vol. 10. No. 7. P. 1—9.

Flora D. B., Curran P. J. (2004) An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data. *Psychological Methods*. Vol. 9. No. 4. P. 466—491. https://doi.org/10.1037/1082-989X.9.4.466.

Knapp T. R. (1990) Treating Ordinal Scales as Interval Scales: An Attempt To Resolve the Controversy. *Nursing Research*. Vol. 39. No. 2. P. 121—123.

Kuzon W. M., Urbanchek M. G., McCabe S. (1996) The Seven Deadly Sins of Statistical Analysis. *Annals of Plastic Surgery*. Vol. 37. No. 3. P. 265—272.

Meulman J. J., Kooij A. J. van der, Heiser W. J. (2004) Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data. Handbook of Quantitative Methodology for the Social Sciences. Ed. by D. Kaplan. Thousand Oaks, Calif.: Sage Publications.

O'Brien R.M. (1979) The Use of Pearson's with Ordinal Data. *American Sociological Review*. Vol. 44. P. 851—857. https://doi.org/10.2307/2094532.

Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T. (1992) Numerical Recipes in FORTRAN. In: The Art of Scientific Computing. Cambridge: Cambridge University Press. P. 2—9.

Rigdon E. E., Ferguson C. E. (1991) The Performance of the Polychoric Correlation Coefficient and Selected Fitting Functions in Confirmatory Factor Analysis with Ordinal Data. *Journal of Marketing Research*. Vol. 28. No. 4. P. 491—497. https://doi.org/10.2307/3172790.

Suits D. B. (1957) Use of Dummy Variables in Regression Equations. *Journal of the American Statistical Association*. Vol. 52. No. 280. P. 548—551. https://doi.org/10.2307/2281705.

Vermunt J. K., Magidson J. (2005) Factor Analysis with Categorical Indicators: A Comparison between Traditional and Latent Class Approaches. In: Quantitative methodology series. New developments in categorial data analysis for the social and behavioral sciences / Ed. by L. A. van der Ark, M. A. Croon, & K. Sijtsma. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. P. 41—62.