

Анализ факторов качества жизни населения России и Европы в рамках метода обобщенных главных компонент

© 2019 г. М.И. Волкова

ЦЭМИ РАН

E-mail: frauwulf@gmail.com

Поступила в редакцию 07.08.2018 г.

Работа выполнена при финансовой поддержке РФФИ (проект 17-18-01080).

По ряду частотных характеристик для набора переменных отдельно для регионов и городов России и стран Европы построены обобщенные главные компоненты для временного интервала 2011—2015 гг. в рамках метода STATIS. Используются данные RLMS и EUROBAROMETER. Отбор критериев для анализа проведен в соответствии с процедурой оценивания связи между порядковыми переменными. Рассмотрены оценки уровня материального благосостояния, статуса респондентов, а также ряда критериев социальной напряженности (частота употребления алкоголя, закредитованность населения, оценка проблемы иммиграции, безработицы). Для этого были использованы полученные непосредственно от респондентов оценки. В рамках метода STATIS удалось выявить группы взаимосвязанных показателей, формирующих максимальный вклад в компромиссное пространство. Компромиссное пространство строится по всем временным интервалам и всей совокупности рассматриваемых признаков. Предложены и апробированы способы вычисления элементов компромиссной (обобщенной) матрицы с учетом критерия максимальной информативности. Внутри компромиссного пространства определены группы населенных пунктов и стран. В качестве критериев разделения на группы использованы полученные в результате сингулярного разложения компромиссной матрицы элементы матрицы счетов. Научная новизна исследования состоит в применении метода обобщенных главных компонент для данных субъективного характера, имеющих вид порядковых переменных. Сопоставляются данные, полученные в ходе опросов населения регионов России и стран Европы.

Ключевые слова: качество жизни, удовлетворенность жизнью, многомерный статистический анализ, обобщенные главные компоненты, RLMS, EUROBAROMETER.

Классификация JEL: C10, C19, C39, C80, D10, I30, I31.

DOI: 10.31857/S042473880004678-4

ВВЕДЕНИЕ

Большие объемы данных — явление, привычное для различных социально-экономических исследований. Однако число инструментов, позволяющих отслеживать динамику и выявлять слабые места в рядах наблюдений для фиксированного числа объектов в виде матриц размера $n \times p$, ограничено.

Особенное значение придается подобным методам в рамках исследования социально-экономических и социально-психологических категорий: “качество жизни” и “удовлетворенность жизнью” (Berenger, Verdier-Chouchane, 2007). Отметим, что основной характеристикой категории “Качество жизни” является отсутствие единого, общепризнанного ее определения, несмотря на большое число методологических концепций ее оценивания (Айвазян и др., 1989). Согласно определению, данному в (Айвазян и др., 1989), “качество жизни” — уникальная категория, являющаяся одновременно *синтетической*, объединяющей различные категории, сферы жизнедеятельности, и *латентной*, непосредственно измерить которую не представляется возможным. Однако если ставится задача субъективного определения качества жизни, то в некоторой степени возможна интеграция ряда оценок, полученных непосредственно от индивидов (Волкова, 2017).

В работе используется методология обобщенных главных компонент, инструментарий для оценивания которых был впервые предложен Бернардом Флури (Flury, 1986). Однако мы ориентиру-

емся на несколько иную концепцию, а именно — на *обобщение метода главных компонент* для динамического ряда (Structuring Three-way Data Sets in Statistics, STATIS) (Escoufier, 1980; Lavit, 1988; Lavit et al., 1994; Abdi, Valentin, 2007). Метод позволяет сравнивать между собой наборы данных за разные интервалы времени, выявлять сходства и различия между отрезками времени, объектами наблюдения и выявлять те из них, которые в наибольшей степени формируют различные проблемы в структуре данных. Положим, исходные данные представляют собой таблицы значений p исходных признаков ($j = 1, \dots, p$) для i наблюдений ($i = 1, \dots, n$) по ряду лет $t = 1, \dots, k$. К данным предъявляется требование: набор объектов — городов, регионов, стран должен быть постоянен для всего исследуемого временного интервала.

Мы предложили и апробировали варианты определения элементов компромиссной матрицы (ее определение будет дано далее):

- выбор матрицы, соответствующей моменту времени $t = 1$, для отслеживания динамики относительно первой точки отсчета;
- определение компромиссной матрицы через усреднение имеющихся наборов данных по числу рассматриваемых временных интервалов;
- вычисление элементов компромиссной матрицы путем взвешивания значений исходных переменных в стандартизованном виде с помощью нормированных весовых коэффициентов. В качестве указанных коэффициентов выступают нормированные компоненты собственного вектора корреляционной матрицы (по временным отрезкам), соответствующего ее максимальному собственному значению.

Основной критерий выбора метода расчета элементов компромиссной матрицы — максимальная доля общей вариации исходных признаков, которую объясняет ее первая главная компонента.

ДАННЫЕ

Рассматриваются результаты опроса респондентов из 35 российских населенных пунктов (Приложение, п. 1) и 34 государства Европы за 2011—2015 гг. по базам RLMS и EUROBAROMETER; расчеты проводились в среде Matlab R2017.

В итоговый перечень переменных вошли те признаки, которые характеризуют качество работы, материального благосостояния, здоровья, оценки уровня закрежденности российских респондентов (Зубаревич, 2005). Также рассматривались комплексные критерии, позволяющие оценивать качество жизни в целом (Ратникова, Фурманов, 2014). Для случая европейских стран дополнительно рассматривались оценки миграционной активности, экономической ситуации, качества образования и здравоохранения в странах проживания респондентов (Johansson, 2002).

Перечень переменных был сформирован неслучайно. Признаки, отражающие отношение респондентов к какому-либо вопросу в обеих базах, — порядковые или бинарные переменные. По ним был проведен анализ таблиц сопряженности, а именно проверка взаимосвязи между рассматриваемыми порядковыми переменными с помощью γ -критерия Гудмана — Краскела (Goodman, Kruskal, 1954, 1959, 1963; Кендалл, Стьюарт, 1973).

Процедура расчета значений критерия основана на ранжировании откликов двух порядковых переменных (Vollen, Barb, 1981), — так с его помощью решается задача определения силы согласованности двух переменных (a , b), относящихся к одному набору объектов. В нашем случае таких наборов два: перечень населенных пунктов по выборке RLMS и набор стран по выборке EUROBAROMETER (O'Brien, 1979). Диапазон значений для γ меняется от -1 до $+1$. Причем при $\gamma = -1$ между переменными наблюдается полная обратная взаимосвязь, а при $\gamma = +1$ — полная прямая взаимосвязь. Если переменные независимы, то значение γ -критерия равно нулю (Гржибовский, 2008).

Основанием для включения переменных в список являлось попадание значения γ -критерия для пары переменных в один из двух интервалов: $-1 \leq \gamma \leq -0,3$ или $0,3 \leq \gamma \leq 1$, в соответствии со смысловой нагрузкой объясняющей переменной.

Исходные данные — переменные с бинарным $x_{ir}^{l(k)} = \begin{cases} 1(k_1); \\ 0(k_2) \end{cases}$ или множественным откликом

$$x_{ir}^{l(k)} = \begin{cases} 1(k_1); \\ \dots \\ g(k_g), \end{cases}$$

где i — номер населенного пункта, r — номер респондента, l — номер вопроса с числом градаций k .

Все исходные переменные были преобразованы в частотные характеристики: $x_{ig}^{l(k)} = \frac{1}{n} \sum_{i=1}^n x_{ir}^{l(k)} \times 100\%$.

Учитывались два нюанса. Во-первых, рассматривались только валидные частоты, т.е. отбрасывались пропущенные значения или пользовательские миссинги (отсутствие ответа на поставленный вопрос): “затрудняюсь ответить”, “отказываюсь отвечать”. Во-вторых, объектом исследования становились не респонденты, а населенные пункты, что достигалось за счет усреднения суммарного отклика для каждой градации по числу респондентов в конкретном населенном пункте (стране). Это было сделано с целью проведения динамического сопоставления в течение ряда лет. Рассматривались только некоторые градации из всех относящихся к конкретной переменной. Так, если у переменной — нечетное число градаций, то выделялись только те, которые находятся выше средней (их частоты суммировались). В случае если у переменной — две градации (1 и 0), то использовалась только градация 1. Опишем методику отбора на конкретных примерах из базы RLMS.

Пример 1. Переменная $x^{5(5)}$. “Насколько вы удовлетворены своим материальным положением? Возможные градации: 1 — полностью удовлетворен; 2 — в принципе удовлетворен; 3 — и да, и нет; 4 — не очень удовлетворен; 5 — совсем не удовлетворен”.

Выделяя и суммируя частоты для *первых двух* градаций, мы имеем дело с теми респондентами, которые удовлетворены своим материальным положением.

Пример 2. Переменная $x^{11(2)}$. “Есть ли у вашей семьи неоплаченные долги по кредитам? Возможные градации: 0 — нет; 1 — да”.

Для дальнейшего анализа этой переменной были использованы только частоты для ответа 1.

Разная смысловая нагрузка переменных обуславливает необходимость преобразования исходной статистической информации (была применена процедура стандартизации).

МЕТОДОЛОГИЯ

Опишем методологию STATIS в виде итеративной процедуры.

Итерация 1. Вычисление матричного произведения: $S_t = X_t X_t^T$, $S_t (n \times n)$, где $X_t (n \times p)$ — матрица стандартизованных значений для года t , $t = 1, \dots, k$ (Rivadeneira et al., 2016), а $X_t^T (p \times n)$ — транспонированная матрица стандартизованных значений.

Итерация 2. Определение скалярного произведения: $\langle S_t, S_{t'} \rangle_{H-S} = \text{trace}(S_t^T S_{t'})$ (Rivadeneira et al., 2016), где $S_{t'} = X_{t'} X_{t'}^T$ — произведение матриц стандартизованных значений за тот год t' , который сопоставляется с текущим годом, $t' = 1, \dots, k-1$.

Итерация 3. Вычисление коэффициентов корреляции r :

$$r_{t,t'} = \frac{\text{trace}(S_t^T S_{t'})}{\sqrt{\text{trace}(S_t^T S_t) \text{trace}(S_{t'}^T S_{t'})}}$$

(Rivadeneira et al., 2016). Матрица $C (k \times k)$ — корреляционная матрица, анализируя элементы которой можно определить пары лет, характеризующиеся большим или меньшим (статистическим) сходством.

Итерация 4. Спектральное разложение корреляционной матрицы C : $C = UVU^{-1}$, где $U^{-1}U = I$. U — матрица, состоящая из собственных векторов матрицы C ; U^{-1} — матрица, обратная ей, а V — диагональная матрица, на диагонали которой — собственные значения корреляционной матрицы C . Могут быть вычислены компоненты матрицы G (координаты каждого набора данных на плоскости главных компонент матрицы C): $G = UV^{1/2}$ ($n \times n$) (Abdi, Valentin, 2007).

Первый столбец матрицы U — собственный вектор, соответствующий максимальному собственному значению матрицы C . Тогда $a_t = u_t / \sum_{t=1}^K u_t$ — нормированные весовые коэффициенты.

Итерация 5. Определение элементов **компромиссной (обобщенной) матрицы** S_{com} ($n \times n$), компоненты которой рассчитываются как линейная комбинация $S_{com} = \sum_{t=1}^K a_t S_t$ объектов, взвешенных с помощью коэффициентов, рассчитанных на предыдущем шаге (Rivadeneira et al., 2016). Для случаев если этот вариант построения дает невысокую информативность, нами были предложены иные методы. Среди них:

- вычисление среднего значения по всем матрицам $S_t = X_t X_t^T$;
- выделение в качестве компромиссной матрицы одной из матриц $S_t = X_t X_t^T$ и другие.

Итерация 6. Спектральное разложение компромиссной матрицы: $S_{com} = QLQ^{-1}$ (Rivadeneira et al., 2016). В результате определяются компоненты матрицы счетов и матрицы нагрузки.

АНАЛИЗ ДАННЫХ RLMS

С помощью расчета значений γ -критерия Гудмана—Краскела был отобран ряд переменных (указан вместе с теми градациями, которые были использованы).

Перечень переменных (RLMS).

$x^{1(2)}$: “Насколько вы удовлетворены своей работой в целом?” (уд_работой) — k_1 — полностью удовлетворен; k_2 — в принципе удовлетворен.

$x^{2(2)}$: “Насколько вы удовлетворены условиями своего труда?” (уд_усл_труда) — k_1 — полностью удовлетворен; k_2 — в принципе удовлетворен.

$x^{3(2)}$: “Насколько вы удовлетворены своей заработной платой?” (уд_зарплатой) — k_1 — полностью удовлетворен; k_2 — в принципе удовлетворен.

$x^{4(2)}$: “Насколько вы удовлетворены возможностями своего профессионального роста?” (проф_рост) — k_1 — полностью удовлетворен; k_2 — в принципе удовлетворен.

$x^{5(2)}$: “Насколько вы удовлетворены своим материальным положением в настоящее время?” (уд_мат_положением) — k_1 — полностью удовлетворен; k_2 — в принципе удовлетворен.

$x^{6(2)}$: “Представьте себе лестницу, где на первой ступени — богатые, а на девятой — нищие. На какой находитесь лично вы в настоящее время (изначально таких ступеней девять)?” (богатство) — k_1 — первая ступень (высшая); k_2 — вторая ступень; k_3 — третья ступень; k_4 — четвертая ступень.

$x^{7(2)}$: “Представьте себе лестницу, где на первой ступени находятся те, кого больше всего уважают, а на девятой — те, кого совсем не уважают. На какой находитесь лично вы в настоящее время (изначально таких ступеней девять)?” (уважение) — k_1 — первая ступень (высшая); k_2 — вторая ступень; k_3 — третья ступень; k_4 — четвертая ступень.

$x^{8(2)}$: “Как вы оцениваете свое здоровье в настоящее время?” (здоровье) — k_1 — очень хорошее; k_2 — хорошее.

$x^{9(2)}$: “Курите ли вы в настоящее время?” (курение) — 1 — да.

$x^{10(2)}$: “Вы употребляете хотя бы иногда алкогольные напитки, включая пиво?” (алкоголь) — 1 — да.

$x^{11(2)}$: “На сегодняшний день у вашей семьи есть какие-либо долги по кредитам?” (кредиты) — 1 — да.

В табл. 1 приведена матрица C коэффициентов корреляции по российским данным.

Таблица 1. Матрица C коэффициентов корреляции (RLMS)

Год	2011	2012	2013	2014	2015
2011	1	0,71	0,60	0,60	0,38
2012	0,71	1	0,66	0,67	0,37
2013	0,60	0,66	1	0,63	0,38
2014	0,60	0,67	0,63	1	0,56
2015	0,38	0,37	0,38	0,56	1

Сходство между соседними интервалами логически объясняется их близостью на шкале времени. Тогда как значительный разрыв между начальной и конечной точкой — результат изменения субъективных оценок, обусловленных в том числе снижением уровня удовлетворенности респондентов материальным положением, работой. Причина такого снижения кроется в первую очередь в последствиях экономического кризиса 2014 г. В результате сингулярного разложения $C = UVU^{-1}$ матрицы C были определены ее собственные значения, находящиеся на диагонали матрицы V :

$$V = \begin{bmatrix} 3,25 & 0 & 0 & 0 & 0 \\ 0 & 0,74 & 0 & 0 & 0 \\ 0 & 0 & 0,40 & 0 & 0 \\ 0 & 0 & 0 & 0,33 & 0 \\ 0 & 0 & 0 & 0 & 0,26 \end{bmatrix} —$$

и собственные векторы — матрица U :

$$U = \begin{bmatrix} 0,458 & -0,295 & -0,634 & 0,361 & 0,412 \\ 0,478 & -0,303 & -0,172 & -0,307 & -0,745 \\ 0,455 & -0,229 & 0,720 & 0,148 & 0,478 \\ 0,479 & 0,171 & 0,184 & -0,690 & 0,479 \\ 0,352 & 0,859 & -0,123 & 0,278 & -0,209 \end{bmatrix}.$$

Определение компонент компромиссной матрицы в случае российских данных производилось двумя способами.

Построение линейной комбинации $S_{com} = \sum_{t=1}^K a_t S_t$, $S_{com}(n \times n)$, где $a_t = u_t / \sum_{t=1}^K u_t$ — нормированные компоненты собственного вектора корреляционной матрицы C , соответствующего ее максимальному собственному значению. Отсюда получаем $a_t = [0,206 \ 0,215 \ 0,205 \ 0,215 \ 0,159]^T$.

Усреднение исходных матриц по числу рассматриваемых K периодов времени:

$$S_{com} = \left(\sum_{t=1}^K S_t \right) / K.$$

По результатам определения объясняющей силы двух компромиссных матриц для дальнейшего анализа был выбран второй вариант. В качестве иллюстрации полученных нами результатов приведем первые пять собственных значений матрицы. Соответствующие им пять первых главных компонент компромиссной матрицы вместе объясняют более 70% общей вариации признаков (табл. 2). Тогда как первый вариант определения элементов компромиссной матрицы дает меньшую информативность — пять первых главных компонент объясняют лишь 65% общей вариации.

Для анализа структуры данных и проекции объектов на плоскость осей (главных компонент) компромиссной матрицы к ней, согласно итерационной процедуре, было проведено спектральное разложение $S_{com} = QLQ^{-1}$. По результатам спектрального разложения и дальнейшего анализа

Таблица 2. Собственные значения компромиссной матрицы (сегмент матрицы L) и доля общей вариации признаков (для семи первых главных компонент), вариант с выбором в качестве компромиссной матрицы среднего значения за 5 лет

Собственное значение	Накопленная доля общей вариации признаков, %
$\lambda_1 = 65,35$	28,9
$\lambda_2 = 40,26$	46,5
$\lambda_3 = 25,03$	57,5
$\lambda_4 = 17,69$	65,3
$\lambda_5 = 16,25$	72,5
...	...
$\sum_l \lambda_l = 227,03$	100

матрицы нагрузок определились группы признаков (максимально коррелированные друг с другом и с осями компромиссного пространства):

- оценка материальной обеспеченности (причем уровень удовлетворенности материальным положением тесно связан с субъективной оценкой здоровья);
- удовлетворенность работой и условиями труда;
- наличие вредных привычек, статус в обществе (рис. 1).

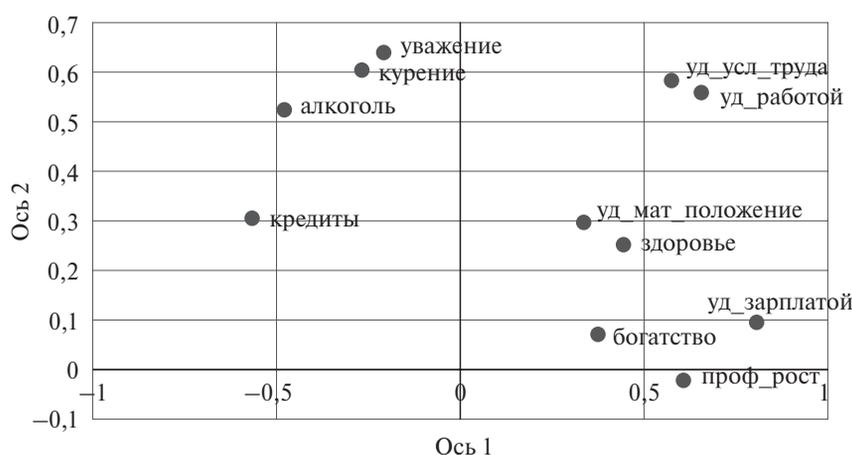


Рис. 1. График нагрузок

Рассчитаны и построены проекции объектов (населенных пунктов) на плоскости первых двух главных компонент компромиссной матрицы (рис. 2) (Vermunt, Magidson, 2005). Для этого определены элементы матрицы счетов: $F_{com} = QL^{1/2}$, размерностью $(n \times n)$, где Q — матрица, состоящая из собственных векторов компромиссной матрицы, а L — диагональная матрица, на диагонали которой — собственные значения компромиссной матрицы, т.е. $L^{1/2}$ — квадратный корень матрицы L . Каждая строка матрицы F_{com} — объект наблюдения (населенный пункт), а каждый столбец — компонента компромиссного пространства.

Произошло ожидаемое разделение всей совокупности объектов на группы. Так, в верхнем правом квадранте — наиболее благополучные объекты. Для них характерны высокие оценки материальной удовлетворенности, возможностей профессионального роста, условий труда. Помимо этого, здесь отмечается низкая степень кредитной задолженности, высокие оценки здоровья и статусные оценки (уважение со стороны окружающих).

Одни из наиболее отстающих — Вольск, Курьинский район (Алтайский край), Руднянский район (Волгоградская область). Здесь отмечаются низкие оценки труда и благосостояния, высокий уровень закредитованности населения (Якубов, 2009).

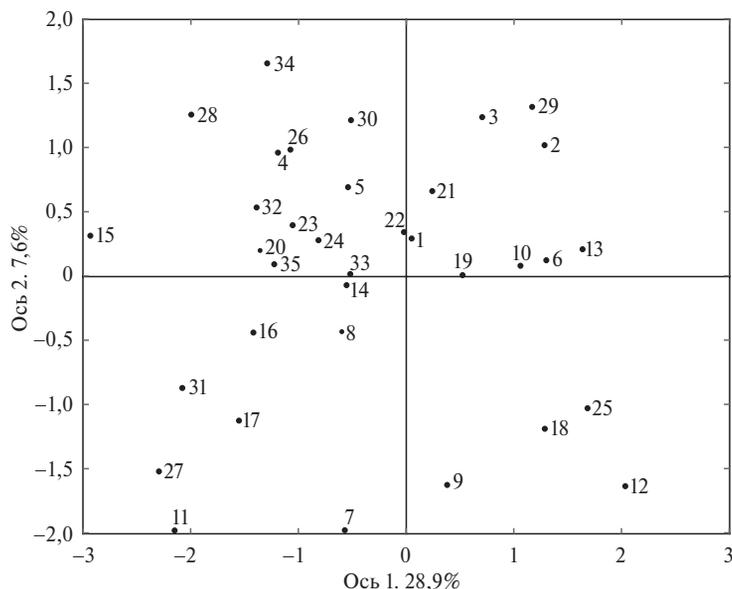


Рис. 2. Проекция объектов на оси (1; 2) компромиссного пространства

Расшифровка: 1 — Санкт-Петербург; 2 — Москва; 3 — Московская область; 4 — Сыктывкар, Коми; 5 — Волосовский район, Ленинградская область; 6 — Смоленск; 7 — Ржев, Тверская область; 8 — Тула; 9 — Куйбышевский район, Калужская область; 10 — Нижний Новгород; 11 — Шумерлинский район, Республика Чувашия; 12 — Земетченский район, Пензенская область; 13 — Уварово, Тамбовская область; 14 — Казань; 15 — Саратов; 16 — Вольск, Саратовская область; 17 — Руднянский район, Волгоградская область; 18 — Батайск, Ростовская область; 19 — Краснодар; 20 — Георгиевск, Ставропольский край; 21 — Кушевский район, Краснодарский край; 22 — Челябинск; 23 — Курган; 24 — Глазов, Удмуртия; 25 — Орск, Оренбургская область; 26 — Соликамск, Пермский край; 27 — Красноармейск, Челябинская область; 28 — Томск; 29 — Сургутский район, Тюменская область; 30 — Бийск, Алтайский край; 31 — Курьинский район, Алтайский край; 32 — Красноярск; 33 — Владивосток; 34 — Назарово, Красноярский край; 35 — Тамбовский район, Амурская область.

Для проверки устойчивости модели из анализа последовательно исключались исходные матрицы (всего было проведено пять вариантов исключения, начиная с 2011 г. и заканчивая 2015 г.). Эти процедуры не повлияли на результаты, информативность компромиссной матрицы не снижалась.

СТРАНЫ ЕВРОПЫ

Для динамического анализа качества жизни в контексте субъективных оценок были отобраны 34 страны (Приложение, п. 2), среди которых есть как страны—члены ЕС, так и государства — кандидаты на вступление в ЕС¹.

Информационная база — стандартный EUROBAROMETER за 2011—2015 гг.²

Из всей совокупности переменных были отобраны девять признаков, имеющие максимальную связь с переменной “Удовлетворенность жизнью” (степень тесноты связи определялась путем расчета γ -критерия Гудмана — Краскела) (Dehley et al., 2002).

Преобразование переменных производилось по той же схеме, которая применялась для российских населенных пунктов.

Оценивались переменные (указаны рассматриваемые градации):

$x_{[EB]}^{(1)}$: “Как вы оцениваете свою работу?” — k_1 : очень хорошо;

¹ Лиссабонский договор, изменяющий “Договор о Европейском союзе” и “Договор об учреждении Европейского сообщества” (Лиссабон, 13 декабря 2007 г.). Текст на русском языке см.: Портал “Право Европейского союза”: <http://eulaw.ru/treaties/lisbon/>; “Договор о Европейском союзе” (подписан в Маастрихте 7.02.1992. Текст на русском языке: <http://ppt.ru/newtext.phtml?id=25293>).

² <https://www.gesis.org>. Leibniz-Institut für Socialwissenschaften (информационный ресурс, база EUROBAROMETER).

- $x_{[EB]}^{2(1)}$: “Как вы оцениваете финансовое положение семьи?” — k_1 : очень хорошо;
- $x_{[EB]}^{3(1)}$: “Как вы оцениваете ситуацию в занятости в стране?” — k_1 : очень хорошо;
- $x_{[EB]}^{4(1)}$: “Вас беспокоит проблема инфляции?” — k_1 : да;
- $x_{[EB]}^{5(1)}$: “Вас беспокоит проблема безработицы?” — k_1 : да;
- $x_{[EB]}^{6(1)}$: “Вас беспокоит проблема иммиграции?” — k_1 : да;
- $x_{[EB]}^{7(1)}$: “Вас беспокоит состояние системы здравоохранения?” — k_1 : да;
- $x_{[EB]}^{8(1)}$: “Вас беспокоит состояние системы образования?” — k_1 : да;
- $x_{[EB]}^{9(1)}$: “Вас беспокоит состояние окружающей среды?” — k_1 : да.

В соответствии с методологией анализа матричнозначных временных рядов были рассчитаны компоненты матрицы $C_{[EB]}$ (для унифицированного изложения результатов оставлены те же обозначения, что и в предыдущем случае, но добавлен нижний индекс $[EB]$ (сокращение от EUROBAROMETER). Компоненты корреляционной $(k \times k)$ -матрицы $C_{[EB]}$

$$r_{t,t'[EB]} = \frac{\text{trace}(S_t^T [EB] S_{t'[EB]})}{\sqrt{\text{trace}(S_t^T [EB] S_{t[EB]}) \text{trace}(S_{t'[EB]}^T S_{t'[EB]})}}$$

(табл. 3). Здесь $S_{t[EB]} = X_{t[EB]} X_{t[EB]}^T$ для каждого периода $t = 1, \dots, k$, а также $S_{t'[EB]} = X_{t'[EB]} X_{t'[EB]}^T$ для периода $t' = 1, \dots, k - 1$; $X_{t[EB]}(n \times p)$ и $X_{t'[EB]}^T(n \times p)$ — исходная и транспонированная матрицы стандартизованных значений рассматриваемых признаков.

Таблица 3. Корреляционная матрица $C_{[EB]}$ (по данным EUROBAROMETER)

Год	2011	2012	2013	2014	2015
2011	1	0,86	0,82	0,78	0,76
2012	0,86	1	0,93	0,90	0,81
2013	0,82	0,93	1	0,92	0,84
2014	0,78	0,90	0,92	1	0,91
2015	0,76	0,81	0,84	0,91	1

Сопоставив элементы матрицы C (по данным RLMS) и $C_{[EB]}$ (по данным EUROBAROMETER), мы отмечаем более сильную связь между отрезками времени во втором случае. Кроме того, определяется большее сходство между соседними отрезками. Минимальное значение коэффициента корреляции между 2011 и 2015 г. — с течением времени между ними произошло некоторое число преобразований, самое заметное из которых — изменение отношения респондентов к иммиграции. Особенно это касается тех стран, куда были направлены наибольшие миграционные потоки из регионов Северной Африки и Ближнего Востока, — Германия, Италия, Швеция (Суворова, 2018).

Определены собственные значения матрицы $C_{[EB]}$:

$$V_{[EB]} = \begin{bmatrix} 4,41 & 0 & 0 & 0 & 0 \\ 0 & 0,29 & 0 & 0 & 0 \\ 0 & 0 & 0,17 & 0 & 0 \\ 0 & 0 & 0 & 0,07 & 0 \\ 0 & 0 & 0 & 0 & 0,05 \end{bmatrix}.$$

Учитывая разницу между первым и вторым собственными значениями матрицы $C_{[EB]}$, можно ожидать большей информативности компромиссного пространства, построенного по европейским данным. Матрица собственных векторов корреляционной матрицы $C_{[EB]}$ имеет вид:

$$U_{[EB]} = \begin{bmatrix} 0,438 & -0,547 & -0,572 & 0,019 & -0,426 \\ 0,458 & -0,360 & 0,117 & 0,188 & 0,782 \\ 0,457 & -0,021 & 0,476 & -0,731 & -0,173 \\ 0,456 & 0,236 & 0,424 & 0,644 & -0,376 \\ 0,427 & 0,718 & -0,502 & -0,125 & 0,186 \end{bmatrix}.$$

Исходя из значений полученных нормированных весовых коэффициентов $a_{t[EB]} = [0,196 \ 0,205 \ 0,205 \ 0,204 \ 0,190]^T$, наименьший вклад в компромиссное пространство имеют матрицы за 2011—2015 гг. Компоненты компромиссной матрицы для европейских стран определены в виде $S_{com[EB]} = \sum_{t=1}^K a_{t[EB]} S_{t[EB]}$, поскольку первая главная компонента компромиссной матрицы объясняет в этом случае 59% общей вариации признаков.

Построена проекция 34 объектов (стран) на плоскость первых двух главных компонент компромиссной матрицы (рис. 3).

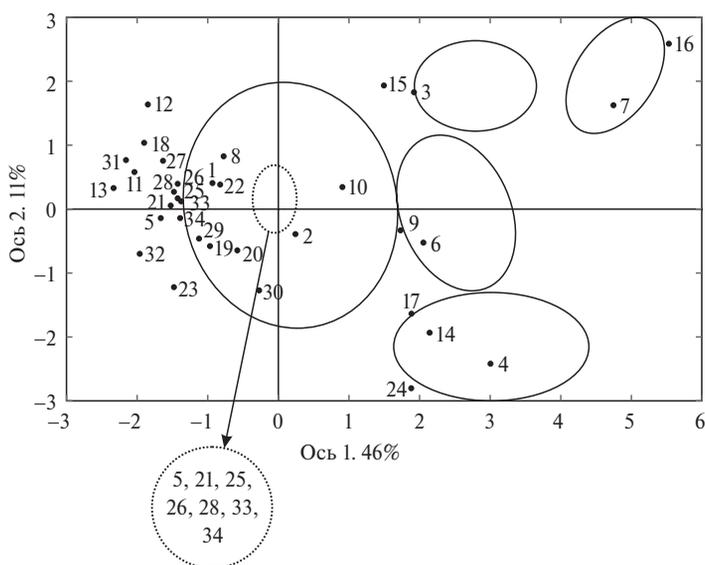


Рис. 3. Проекция стран на оси (1; 2) компромиссного пространства

Расшифровка: 1 — Франция; 2 — Бельгия; 3 — Голландия; 4 — Германия (Запад); 5 — Италия; 6 — Люксембург; 7 — Дания; 8 — Ирландия; 9 — Великобритания; 10 — Северная Ирландия; 11 — Греция; 12 — Испания; 13 — Португалия; 14 — Германия (Восток); 15 — Финляндия; 16 — Швеция; 17 — Австрия; 18 — Кипр; 19 — Чехия; 20 — Эстония; 21 — Венгрия; 22 — Латвия; 23 — Литва; 24 — Мальта; 25 — Польша; 26 — Словакия; 27 — Словения; 28 — Болгария; 29 — Румыния; 30 — Турция; 31 — Хорватия; 32 — Кипр (Турецкая часть); 33 — Македония; 34 — Черногория.

Вся совокупность стран очень четко разделилась на группы. С одной стороны, богатые страны с развитой экономикой, население которых, по большей части, довольно своим материальным положением, работой, высоко оценивает состояние систем образования и здравоохранения, а также выражает обеспокоенность миграционными процессами. С другой — страны юга и центра Европы. Особенно выделяются Испания, Кипр, Хорватия, Португалия, Греция (точка с номером 11 на рис. 3). Для них характерны высокий уровень безработицы, невысокая (субъективная) материальная обеспеченность домашних хозяйств и множество других, сопряженных с перечисленными, проблем. Последовательное исключение исходных матриц из анализа не нарушило устойчивости модели.

ЗАКЛЮЧЕНИЕ

Для анализа матричнозначных временных рядов на диапазоне 2011—2015 гг. к данным субъективного характера был применен ряд матричных преобразований в рамках метода STATIS. Опре-

делена теснота связи между рассматриваемыми временными отрезками для обоих случаев. Для российских данных коэффициенты корреляции между матрицами объект-признак по российским данным ниже, чем в случае анализа стран Европы. Это в первую очередь обусловлено более значимыми изменениями в оценках, полученных от респондентов, проживающих в Российской Федерации. Компромиссные (обобщенные) матрицы получены для двух совокупностей исходных данных. В обоих случаях отмечена устойчивость полученных результатов. Определены группы объектов, расположенных максимально близко относительно друг друга на плоскости первых двух главных компонент обобщенных матриц. Разделение европейских стран на группы более иллюстративно: четко отделены друг от друга группы благополучных и неблагополучных стран. В список самых благополучных попали две страны: Дания и Швеция. Москва, Московская область и город Сургут — тройка лидеров по субъективному качеству жизни среди российских населенных пунктов.

ПРИЛОЖЕНИЕ

1. Перечень населенных пунктов и переменных, вошедших в анализ удовлетворенности жизнью населения (по данным RLMS):

1) Санкт-Петербург; 2) Москва; 3) Московская область; 4) Сыктывкар, Коми; 5) Волосовский район, Ленинградская область; 6) Смоленск; 7) Ржев, Тверская область; 8) Тула; 9) Куйбышевский район, Калужская область; 10) Нижний Новгород; 11) Шумерлинский район, Республика Чувашия; 12) Земетченский район, Пензенская область; 13) Уварово, Тамбовская область; 14) Казань; 15) Саратов; 16) Вольск, Саратовская область; 17) Руднянский район, Волгоградская область; 18) Батайск, Ростовская область; 19) Краснодар; 20) Георгиевск, Ставропольский край; 21) Кушевский район, Краснодарский край; 22) Челябинск; 23) Курган; 24) Глазов, Удмуртия; 25) Орск, Оренбургская область; 26) Соликамск, Пермский край; 27) Красноармейск, Челябинская область; 28) Томск; 29) Сургутский район, Тюменская область; 30) Бийск, Алтайский край; 31) Курьинский район, Алтайский край; 32) Красноярск; 33) Владивосток; 34) Назарово, Красноярский край; 35) Тамбовский район, Амурская область.

2. Перечень населенных пунктов и переменных, вошедших в анализ удовлетворенности жизнью населения (по данным Eurobarometer):

1) Франция; 2) Бельгия; 3) Голландия; 4) Германия (Запад); 5) Германия (Восток) (комментарий: в таком виде, с учетом территориального разделения, приводится информация в системе EUROBAROMETER, поскольку прежде ФРГ и ГДР приписывались разные весовые коэффициенты, которые сейчас не используются); 6) Люксембург; 7) Дания; 8) Ирландия; 9) Великобритания; 10) Северная Ирландия; 11) Греция; 12) Испания; 13) Португалия; 14) Италия; 15) Финляндия; 16) Швеция; 17) Австрия; 18) Кипр; 19) Чехия; 20) Эстония; 21) Венгрия; 22) Латвия; 23) Литва; 24) Мальта; 25) Польша; 26) Словакия; 27) Словения; 28) Болгария; 29) Румыния; 30) Турция; 31) Хорватия; 32) Кипр (Турецкая часть); 33) Македония; 34) Черногория.

СПИСОК ЛИТЕРАТУРЫ

- Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д. (1989). Прикладная статистика: Классификация и снижение размерности. Айвазян С.А. (общ. ред.). М.: Финансы и статистика.
- Волкова М.И. (2017). Выявление факторов удовлетворенности жизнью в России и Европе // *Социальная политика и социология*. № 5. С. 6—15.
- Гржибовский А.М. (2008). Анализ порядковых // *Экология человека*. № 8. С. 56—62.
- Зубаревич Н.В. (2005). Россия регионов: в каком социальном пространстве мы живем? Независимый институт социальной политики. М.: Поматур.
- Кендалл М., Стьюарт А. (1973). Статистические выводы и связи. Т. 2. Колмогорова А.Н. (общ. ред.). М.: Наука, Физматлит.
- Ратникова Т.А., Фурманов К.К. (2014). Анализ панельных данных и данных о длительности состояний. Учебное пособие. НИУ ВШЭ. М.: Изд. дом Высшей школы экономики.

- Суворова В.А.** (2018). Миграционный кризис в Европе: проблемы вынужденной миграции // *Власть*. № 1. P. 176—179.
- Якубов А.Х.** (2009). Цели, критерии и методические подходы к оптимизации сельских поселений в муниципальном районе // *Вестник Челябинского государственного университета*. № 26 (164). С. 104—109.
- Abdi H., Valentin D.** (2007). Multiple Correspondence Analysis. In: Salkind N. (ed.) *“Encyclopedia of Measurement and Statistics”*. Thousand Oaks: Sage. P. 652—658.
- Berenger V., Verdier-Chouchane A.** (2007). Multidimensional Measures of Well-Being: Standard of Living and Quality of Life Across Countries // *World Development*. Vol. 35. No. 7. P. 1259—1276.
- 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>.
- Dehley J., Bpohnke P., Habich R., Zapf W.** (2002). Quality of Life in a European Perspective: The Euromodule as a New Instrument for Comparative Welfare Research // *Social Indicators Research*. Vol. 58. P. 163—176.
- Escoufier Y.** (1980). L’analyse conjointe de plusieurs matrices de données. In: Jolivet M. (ed.) *“Biométrie et Temps. Paris: Société Française de Biométrie”*. P. 59—76. [Электронный ресурс] Режим доступа: https://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/eurobarometer/eb_bibliography/EB_Bibliography.pdf, свободный. Загл. с экрана. Яз. фран. (дата обращения: декабрь 2018 г.).
- Flury B.N.** (1986). Asymptotic Theory for Common Principal Component Analysis // *The Annals of Statistics*. Vol. 14. No. 2. P. 418—430.
- Goodman L., Kruskal W.H.** (1954). Measures of Association for Crossclassification I // *Journal of the American Statistical Association*. Vol. 49. P. 732—764.
- Goodman L., Kruskal W.H.** (1959). Measures of Association for Crossclassification II // *Journal of the American Statistical Association*. Vol. 54. P. 123—163.
- Goodman L., Kruskal W.H.** (1963). Measures of Association for Crossclassification III // *Journal of the American Statistical Association*. Vol. 58, pp. 310—364.
- Johansson S.** (2002). Conceptualizing and Measuring Quality of Life for National Policy // *Social Indicators Research*. Vol. 58. P. 13—32.
- Lavit C.** (1988). Analyse conjointe de tableaux quantitatifs. Paris: Masson.
- Lavit C., Escoufier Y., Sabatier R., Traissac P.** (1994). The ACT (STATIS method) // *Computational Statistics & Data Analysis*. Vol. 18. P. 97—119.
- 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>.
- Rivadeneira F.J., Figueiredo A.M.S., Figueiredo F.O.S., Carvajal S.M., Rivadeneira R.A.** (2016). Analysis of Well-being in OECD Countries Through STATIS Methodology // *HOLOS*. Vol. 7. P. 335—351.
- 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 Categorical Data Analysis for the Social and Behavioral Sciences. Ark L.A. van der, Croon M.A., Sijtsma K. (eds). Mahwah: Lawrence Erlbaum Associates Publishers. P. 41—62.

REFERENCES (with English translation or transliteration)

- Abdi H., Valentin D.** (2007). Multiple Correspondence Analysis. In: Salkind N. (ed.) *“Encyclopedia of Measurement and Statistics”*. Thousand Oaks: Sage, 652—658.
- Aivazian S.A., Buchshtaber V.M., Enyukov I.S., Meshalkin L.D.** (1989). Applied statistics: Classification and dimension reduction. Aivazian S.A. (ed.). Moscow: Finansy i Statistika (in Russian).
- Berenger V., Verdier-Chouchane A.** (2007). Multidimensional Measures of Well-Being: Standard of Living and Quality of Life Across Countries. *World Development*, 35, 7, 1259—1276.
- Bollen K.A., Barb K.H.** (1981). Pearson’s r and Coarsely Categorized Measures. *American Sociological Review*, 46, 232—239. <https://doi.org/10.2307/2094981>.
- Dehley J., Bpohnke P., Habich R., Zapf W.** (2002). Quality of Life in a European Perspective: The Euromodule as a New Instrument for Comparative Welfare Research. *Social Indicators Research*, 58, 163—176.

- Escoufier Y.** (1980). L'analyse conjointe de plusieurs matrices de données. In: Jolivet M. (ed.) "Biométrie et Temps. Paris: Société Française de Biométrie", 59—76. Available at: https://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/eurobarometer/eb_bibliography/EB_Bibliography.pdf (accessed: December 2018).
- Flury B.N.** (1986). Asymptotic Theory for Common Principal Component Analysis. *The Annals of Statistics*, 14, 2, 418—430.
- Goodman L., Kruskal W.H.** (1954). Measures of Association for Crossclassification I. *Journal of the American Statistical Association*, 49, 732—764.
- Goodman L., Kruskal W.H.** (1959). Measures of Association for Crossclassification II. *Journal of the American Statistical Association*, 54, 123—163.
- Goodman L., Kruskal W.H.** (1963). Measures of Association for Crossclassification III. *Journal of the American Statistical Association*, 58, 310—364.
- Grjibovskiy A.M.** (2008). Ordered Data Analysis. *Human Ecology*, 8, 56—62 (in Russian).
- Johansson S.** (2002). Conceptualizing and Measuring Quality of Life for National Policy. *Social Indicators Research*, 58, 13—32.
- Kendall M., Stewart A.** (1973). The Advanced Theory of Statistics. Vol. 2: Inference and Relationship. Gal'chuk L.I., Terekhin A.T. (trans.), Kolmogorov A.N. (ed.). Moscow: Nauka, Fizmatlit (in Russian).
- Lavit C.** (1988). Analyse conjointe de tableaux quantitatifs. Paris: Masson.
- Lavit C., Escoufier Y., Sabatier R., Traissac P.** (1994). The ACT (STATIS Method). *Computational Statistics & Data Analysis*, 18, 97—119.
- O'Brien R.M.** (1979). The Use of Pearson's with Ordinal Data. *American Sociological Review*, 44, 851—857. <https://doi.org/10.2307/2094532>.
- Ratnikova T., Furmanov K.K.** (2014). Panel Data Analysis and Data on the Duration of States. Textbook. Higher School of Economics. Moscow: Izd. dom Vysshei shkoly ekonomiki (in Russian).
- Rivadeneira F.J., Figueiredo A.M.S., Figueiredo F.O.S., Carvajal S.M., Rivadeneira R.A.** (2016). Analysis of Well-being in OECD Countries Through STATIS Methodology. *HOLOS*, 7, 335—351.
- Suvorova V.A.** (2018). Migration crisis in Europe: problems of forced migration. *Power* (IS RAS magazine), 1, 176—179 (in Russian).
- 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 Categorical Data Analysis for the Social and Behavioral Sciences. Ark L.A. van der, Croon M.A., Sijtsma K. (eds). Mahwah: Lawrence Erlbaum Associates Publishers, 41—62.
- Volkova M.I.** (2017). Determining of Factors of Life Satisfaction in Russia and European countries. *Social Policy and Sociology*, 5, 6—15 (in Russian).
- Yakubov A.H.** (2009). Objectives, Criteria and Methodological Approaches to the Optimization of Rural Settlements in the Municipal District. *Herald of Chelyabinsk State Pedagogical University*, 26 (164), 104—109.
- Zubarevich N.V.** (2005). Russia of the Regions: In What Social Space Do We Live? Independent Institute of Social Policy. Moscow: Pomatur (in Russian).

Russian and European Population's Quality of Life Analysis with the Instruments of Common Principal Components (CPC)

© 2019 M.I. Volkova

Central Economics and Mathematics Institute of RAS, Moscow, Russia
E-mail: frauwulf@gmail.com

Received 07.08.2018

This study was carried with financial support by the Russian Science Foundation (project 17-18-01080).

For a number of frequency characteristics of several variables common principal components for period 2011—2015 using the STATIS method is constructed separately for the Russian regions and the European countries. For it, data from RLMS and EUROBAROMETER was used. The selection of criteria

for the analysis was carried within the procedure of assessing γ -criteria for ordinal variables. Estimates of material well-being level, status of respondents, as well as a frequency of alcohol consumption, debt payment difficulties, assessment of immigration and unemployment problem and others are used. Within the framework of STATIS method, it was possible to identify interrelated indicators that form the maximum contribution to the compromise space. The compromise space is built for all the time intervals and features. The methods for calculating the elements of the compromise (generalized) matrix are proposed and tested taking into account the criterion of maximum informativeness. Within the compromise space, groups of the regions (for Russia) and the European countries have been identified. The elements of count matrix obtained as a result of singular decomposition of the compromise matrix are used as criteria of division into groups. The novelty of the research is based on the application of common principal components methodology for subjective data (ordinal variables). The data from the Russian and European surveys are compared.

Keywords: quality of life, life satisfaction, multidimensional statistical analysis, common principal components, RLMS, EUROBAROMETER.

JEL Classification: C10, C19, C39, C80, D10, I30, I31.

DOI: 10.31857/S042473880004678-4