

Факторный анализ показателей нравственного состояния общества в европейских странах

Л. С. Куравский*,

доктор технических наук, профессор, заведующий кафедрой прикладной информатики, декан факультета информационных технологий Московского городского психолого-педагогического университета

А. В. Юревич**,

доктор психологических наук, профессор, член-корреспондент Российской академии наук, заместитель директора Института психологии РАН

П. А. Мармалюк***,

программист кафедры прикладной информатики факультета информационных технологий Московского городского психолого-педагогического университета

Е. Г. Иванова****,

психолог кафедры прикладной информатики факультета информационных технологий Московского городского психолого-педагогического университета

В статье представлен аппарат и результаты факторного анализа динамики развития и взаимодействия характеристик, которые определяют индекс нравственного состояния общества, разработанный Институтом психологии РАН. Примененные для исследования альтернативный вариант конфирматорного факторного анализа и обучаемые сети Маркова позволили выявить ряд закономерностей развития проблемных характеристик, свидетельствующих о правомерности использования указанного индекса. К преимуществам применённого варианта факторного анализа относятся: возможность нахождения оценок свободных параметров модели прямыми (неитерационными) методами, гарантирующими однозначное оптимальное решение, гибкие средства для исследования взаимодействия факторов и применимость для анализа ковариационных и корреляционных матриц произвольной структуры. Сети Маркова, в свою очередь, позволили выявить новые типы показателей, не доступные для традиционных подходов.

Ключевые слова: нравственность, индекс нравственного состояния общества, факторный анализ, сети Маркова.

* kuravsky@yahoo.com

** yurev@orc.ru

*** ykk.mail@gmail.com

**** ivanova_jenny@mail.ru

1. Введение

Нравственность – одна из проблем, к которым психологическая наука обращается нечасто и неохотно по разным причинам, среди них можно выделить следующие:

- психологи часто воспринимают эту проблему как «закрепленную» за философией, в первую очередь, за таким ее разделом, как этика;

- она плохо поддается эмпирическому изучению и «препарированию» в психологических лабораториях;

- естественнонаучная психология, в соответствии с традициями естественной науки, трактует объективность научного исследования как предполагающую абстрагирование от ценностей, что едва ли возможно при изучении таких проблем, как нравственность.

Так или иначе, хотя многие традиционные объекты психологического изучения – ценности, установки и др., имеют непосредственное отношение к морали и нравственности, последняя оказалась вынесенной за пределы основного русла психологической проблематики. Вместе с тем, значительная доля работ в этой области, в том числе и ведущихся в современной отечественной психологии, так или иначе ее затрагивает. В частности, в таком наиболее актуальном для современной России аспекте, как нравственное состояние современного российского общества.

В Институте психологии РАН в русле развиваемой в нем количественной макропсихологии [4; 5] разработан «Индекс нравственного состояния общества» (ИНСО), основанный на интеграции следующих показателей: 1) убийств; 2) беспризорных детей (число на 100 000 жителей); 3) индекс коррупции; 4) индекс Джини, выражающий неравномерность распределения доходов.

Методика вычисления ИНСО такова. Первичные индексы вычисляются как базисные. Значение показателя России за 1996 г. (V_0) считается равным 100 баллам. Нормализация значений показателей (V_i) за более поздние или более ранние годы осуществляется путем их соотношения со значением за 1996 г. (чем выше балл, тем

благополучнее общество по соответствующему параметру). Для нормализации значений показателей социального сиротства, смертности от убийств и индекса Джини, увеличение значений которых указывает на ухудшение состояния в обществе, используется формула

$$Y_i = \frac{V_0}{V_i} \times 100\%.$$

Начиная с 1996 г. интегральный индекс рассчитывается как среднее геометрическое значение четырех первичных индексов:

- смертности от убийств;

- численности выявленных детей, оставшихся без попечительства родителей;

- коррупции;

- неравномерности доходов (индекс Джини).

Ввиду того что до 1996 г. индекс коррупции не рассчитывался, композитный индекс за 1990–1995 гг. строится по трем первичным индексам.

Динамика нравственного состояния российского общества в годы реформ, количественно оцененного таким образом, показана на рис. 1, из которого видно, что нравственное состояние нашего общества ежегодно ухудшалось с 1991 по 1994 гг., затем улучшалось до «дефолтного» 1998 г., после чего вновь ухудшалось до 2002 г., а потом обнаружило тенденцию к ежегодному улучшению (за 2007 и 2008 гг. индекс не рассчитан ввиду того, что статистические данные, лежащие в основе его расчетов, пока отсутствуют). Не подвергая интерпретации выявленную динамику, отметим, что она почти полностью соответствует динамике макропсихологического состояния современного российского общества, оцененной на основе других показателей [6], а также временной развертке его характеристик, вычисляемых социологами (социальные настроения, социальный оптимизм и др.), что говорит о синхронном проявлении подобной динамики в самых различных сферах. Количественная оценка нравственного состояния нашего общества по-

казывает, что в первые годы реформ оно ухудшалось высокими темпами, что свидетельствует о связи данного явления именно с реформами и с сопутствовавшими им событиями. Полученные данные показывают, что в течение последующих после реформ лет обнаруживается нелинейная, «волнообразная» динамика, показатели которой уступают уровню 1990 г. почти в два раза.

Большой интерес для исследования нравственного состояния общества представляет анализ динамики развития и взаимодействия составляющих «Индекс нравственного состояния общества» показателей, а также выявление факторов, которыми обусловлена их изменчивость. Анализ данных европейских стран по указанным выше показателям при помощи традиционных методов, включая регрессионный и корреляционный анализ, не позволил выявить значимые линейные зависимости между выделенными величинами, которые были бы достаточны для их содержательной интерпретации. Поиск нелинейных зависимостей с помощью аппарата нейронных сетей также не дал положительных результатов, что сделало актуальными разработку и применение новых, более сложных методов анализа исследуе-

мых показателей. В данной работе рассматриваются два из них: «Альтернативный вариант конфирматорного факторного анализа» и «Обучаемые сети Маркова», позволившие выявить ряд частных закономерностей развития проблемных характеристик.

2. Факторный анализ

Выявление закономерностей развития и взаимных связей исследуемых характеристик является одной из актуальных задач, возникающих при анализе результатов социологических исследований. Как правило, рассматривается временная динамика наблюдаемых переменных, которые в различных контрольных точках формально рассматриваются как отдельно анализируемые величины. Для решения задачи используются различные методы статистического анализа, наиболее эффективным из которых является симплекс-метод конфирматорного факторного анализа (КФА) [8]. Однако практическое применение этого подхода обусловлено рядом серьезных ограничений, которые зачастую делают его использование невозможным. Кроме того, существенным недостатком традиционного КФА является необходи-

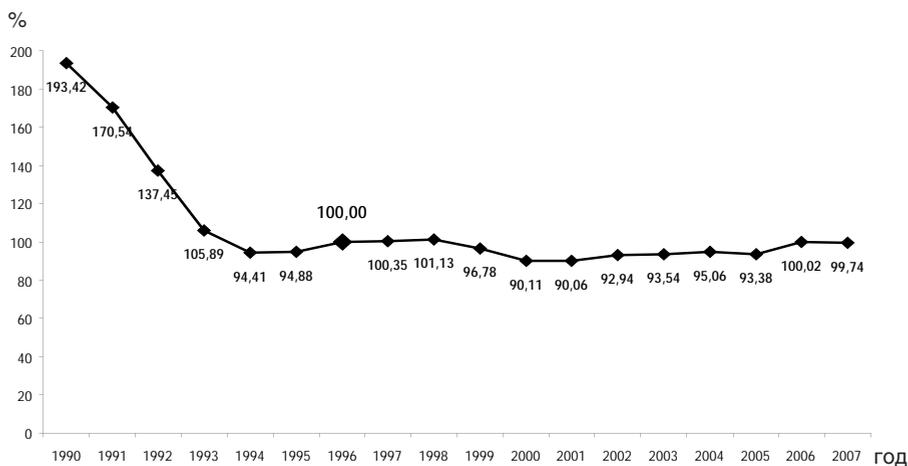


Рис. 1. Динамика нравственного состояния современного российского общества

мость решения трудоемкой задачи многомерной локальной оптимизации для оценки величин свободных параметров модели, что, как правило, не позволяет найти глобальный минимум и приводит к неоднозначности решения.

Чтобы преодолеть указанные проблемы, был разработан новый подход, опирающийся на возможности вейвлет-преобразований и обучаемых факторных структур [10]. К его особенностям и преимуществам относятся:

- возможность нахождения оценок свободных параметров модели прямыми (неитерационными) методами, гарантирующими однозначное оптимальное решение;
- гибкие средства для исследования взаимодействия факторов;
- применимость для анализа ковариационных и корреляционных матриц произвольной структуры.

2.1. Основные этапы факторного анализа

Основные этапы предложенного подхода представлены на рис. 2. Выборки коэффициентов, полученных в результате кратномасштабного вейвлет-анализа

временных рядов значений исследуемого параметра и соответствующих различным периодам наблюдений, в последующем подтверждающем факторном анализе рассматриваются как значения наблюдаемых переменных, которые позволяют выявить временную историю факторных влияний и оценки факторного взаимодействия. Представление данных, полученное с помощью вейвлет-преобразований, дает возможность выявить различия в характеристиках процесса в различных шкалах, что особенно важно при большом числе контрольных моментов времени, связанных с проведением измерений. Идентификация свободных параметров факторной модели (обычно корреляций или ковариаций) выполняется с помощью новой прямой (неитерационной) процедуры, которая опирается на «Метод максимального правдоподобия» (ММП) и является альтернативой традиционному итерационному способу решения задач локальной оптимизации.

Рабочие представления анализируемых процессов формируются с помощью вейвлет-преобразования, которое позволяет выявлять отличия в их характери-

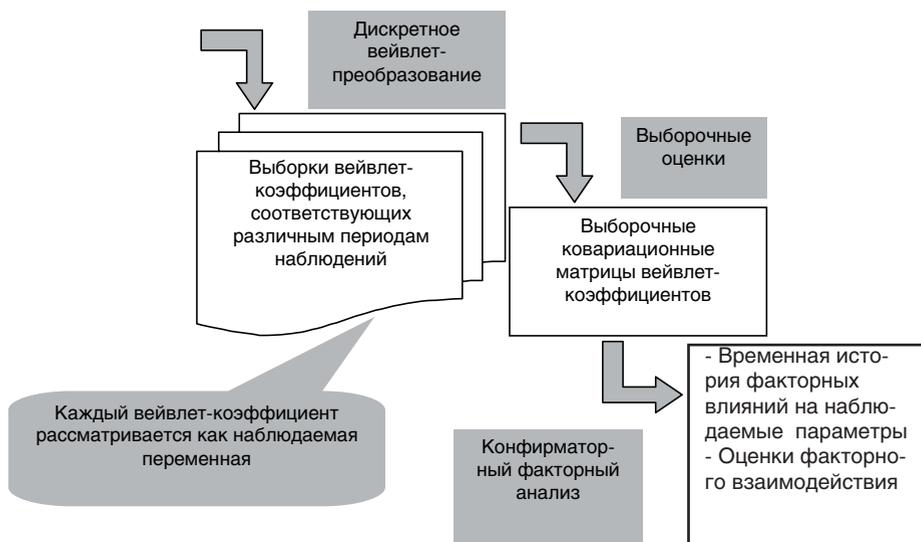


Рис. 2. Основные этапы анализа

ках при использовании различных шкал измерений на всем протяжении интервала наблюдений. Если исследуемый процесс есть функция одной переменной, то его вейвлет-спектр – функция двух аргументов, один из которых характеризует период составляющих компонентов, а другой – смещение вычисляемых показателей вдоль оси времени. Для вычисления вейвлет-спектров используются вейвлеты – особые функции в форме коротких волн («всплесков») с нулевым интегральным значением и локализацией по оси независимой переменной, способные к сдвигу по этой оси и масштабированию.

Вейвлет-анализ имеет очевидные преимущества перед традиционным спектральным анализом, поскольку он обеспечивает корректные результаты в случае нестационарных процессов и содержит более полную информацию о поведении изучаемого объекта. Это сделало данный подход популярным среди исследователей разных специальностей. Далее применялся дискретный вариант этого метода, позво-

ливший представить анализируемые процессы в виде элементов определенного метрического функционального пространства с вейвлет-базисом.

2.2. Альтернативный вариант конфирматорного факторного анализа

Основные компоненты КФА представлены на рис. 3.

КФА предполагает наличие строго определенной факторной модели изучаемого явления. Факторная модель, связывающая латентные и наблюдаемые переменные, формируется, опираясь на знание предметной области. Объектами данного вида анализа являются ковариационные или корреляционные матрицы наблюдаемых переменных. Цель работы – выявить значения параметров модели, которая с приемлемыми ошибками объясняет изменчивость наблюдений.

Для определения свободных параметров модели традиционный КФА предполагает решение трудоемкой задачи многомерной локальной оптимизации. Это приводит к

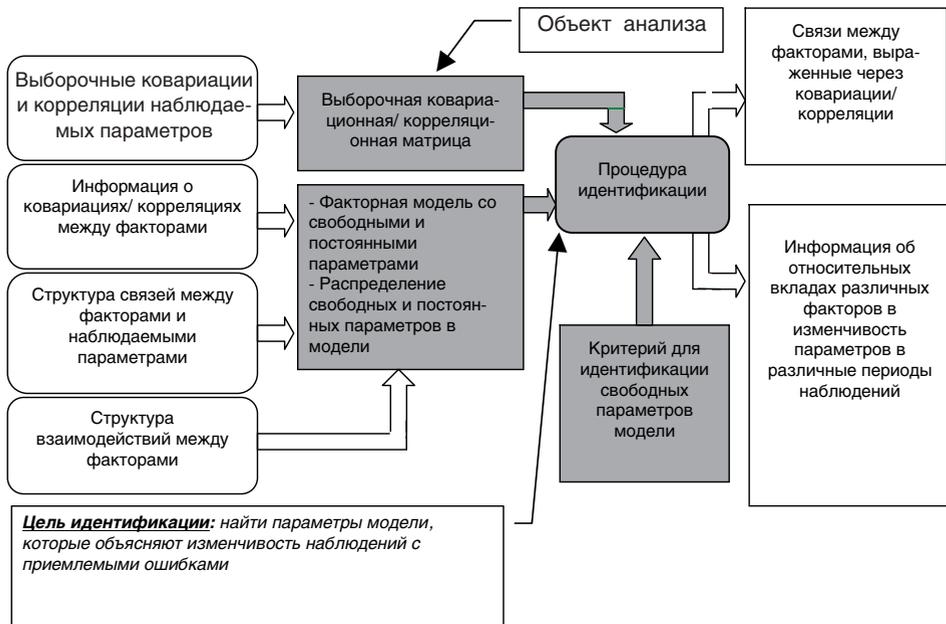


Рис. 3. Основные компоненты КФА

невозможности определения глобального минимума и неоднозначности решения.

Рассматриваемый далее альтернативный вариант КФА позволяет находить единственное оптимальное решение прямыми (неитерационными) методами. Процедура решения включает в себя:

1) составление переопределенной системы алгебраических уравнений при выражении выборочных дисперсий и ковариации через аналогичные факторные показатели¹;

2) решение системы прямым (неитерационным) методом, используя ММП² [2; 12];

3) проверку адекватности полученных моделей наблюдениям при опоре на статистические критерии согласия.

Чтобы избежать составления сложных для решения нелинейных систем уравнений относительно коэффициентов корреляции и факторных нагрузок, используется путевая модель дисперсионных составляющих (рис. 4) [13], в которой факторные нагрузки равны единице. Каждой наблюдаемой дисперсии и ковариации ставится в соответствие алгебраическое уравнение, которое связывает ее выборочную оценку с соответствующей прогнозируемой величиной, выраженной аналитически через неопределенные дисперсии и ковариации латентных переменных [7]. В путевых диаграммах³ для этого, в частности, могут быть использованы правила обхода путей⁴. В результате получается система, число уравнений которой равно числу наблюдаемых дисперсий и ковариаций.

Представим полученную переопределенную систему n уравнений в матричной форме

$$Ax=b,$$

где A – матрица системы, коэффициенты которой определяются факторной моделью; b – вектор-столбец n выборочных дисперсий и ковариаций, определяемых результатами наблюдений; x – вектор-столбец m искоемых дисперсий и ковариаций латентных переменных.

Теперь рассмотрим вектор $\varepsilon=Ax-b$, представляющий полученную методом наименьших квадратов невязку псевдорешения x , переопределенной системы. Полагая в общем случае, что компоненты вектора невязок коррелированы, выразим их невырожденную ковариационную матрицу как σ^2V .

Сделав замены

$$b=V^{1/2}b_0 \text{ и } A=V^{1/2}A_0,$$

где $V=V^{1/2}V^{1/2}$,⁵

перейдем к системе

$$A_0x=b_0,$$

ковариационная матрица вектора невязок $\varepsilon_0=V^{-1/2}\varepsilon$ которой имеет вид σ^2E , где E – единичная матрица.

Если рассматриваемая система невырождена ($rank A = m$, где m – число свободных параметров модели), вектор невязки ε_0 имеет многомерное нормальное распределение, а

$$x_0 = (A_0^T A_0)^{-1} A_0^T b_0 = (A^T V^{-1} A)^{-1} A^T V^{-1} b -$$

псевдорешение, полученное методом наименьших квадратов, это псевдорешение является оценкой максимального правдоподобия, а статистика

$X^2=(b_0-A_0x_0)^T(b_0-A_0x_0)/\sigma^2=(b-Ax_0)^T V^{-1}(b-Ax_0)/\sigma^2$ имеет распределение χ^2 с $n-m$ степенями свободы.

Указанная статистика X^2 позволяет при заданных выше предположениях проверять

¹ Как правило, они являются свободными параметрами модели.

² В другой форме, нежели при традиционном КФА.

³ В путевых диаграммах (рис. 3–7 и 11) овалы (круги) соответствуют латентным переменным, прямоугольники соответствуют наблюдаемым переменным, однонаправленные стрелки соответствуют причинным связям, двунаправленные стрелки соответствуют ковариациям, дисперсиям или корреляциям.

⁴ Обход начинается против причинной связи, затем происходит поворот по ковариационной связи, после чего движение продолжается вдоль причинной связи. При движении запрещается дважды обходить ковариационную связь.

⁵ Для любой симметричной неотрицательно определенной матрицы V (именно к этому классу относятся ковариационные матрицы) существует единственная симметричная неотрицательно определенная матрица $V^{1/2}$, называемая квадратным корнем из V , такая, что $(V^{1/2})^2=V$.

гипотезу о представимости выборочных дисперсий и ковариаций, составляющих вектор \mathbf{b} , дисперсиями и ковариациями латентных переменных, содержащихся в исследуемой модели. Область принятия гипотезы есть $\chi^2 \leq \chi^2_{n-m;\alpha}$, где α есть уровень значимости критерия.

При реализации данного подхода удобно сделать следующие упрощающие предположения, обусловленные особенностями искомого решения:

- компоненты вектора невязок ε являются некоррелированными;
- значения среднеквадратических отклонений различных компонентов вектора невязок ε составляют одну и ту же фиксированную долю (процент) от соответствующих компонентов вектора \mathbf{b} ⁶.

Чтобы обеспечить сопоставимость оценок, указанная доля (процент) подбирается так, чтобы равенство $\chi^2 = \chi^2_{n-m;\alpha}$ выполнялось при $\alpha=0,05$. Для оценки степени допустимости вычисленной характеристики для нее удобно установить разумное критическое значение, например, 0,1. Таким образом, вместо уровня значимости появляется новый показатель – критический процент.

Преимуществом предложенного подхода является то, что он не сводит решение к трудоемкой процедуре многомерной локальной оптимизации, которая не гарантирует нахождение глобальных минимумов,

обеспечивая при этом однозначность результата.

Как и в традиционном КФА, рассматриваемый метод дает возможность строить заключения о статистической значимости различных компонентов модели, используя статистические критерии согласия [3].

2.3. Факторные модели для анализа результатов вейвлет-преобразований

В данной работе применяется модель дисперсионных составляющих. Пример типового варианта таких моделей показан на рис. 4. Состав вейвлет-коэффициентов, используемых при анализе в качестве наблюдаемых переменных, зависит от рассматриваемой прикладной задачи и может меняться. Обычно предполагается, что число моментов времени, в которые производятся наблюдения, является степенью числа 2.

Аналитические выражения для ковариаций и дисперсий вейвлет-коэффициентов W_l линейны:

$$Cov(W_i, W_j) = \sum_k C_{kij}, \quad Var(W_j) = \sum_k V_k + \sum_k \sum_l C_{kl},$$

где k и l – индексы факторов, V^* – дисперсии, C^{**} и C^{***} – ковариации между факторами.

Такие модели дают возможность получать прямые оценки свободных параметров, используя рассмотренный выше альтернативный вариант КФА, и применять данную модель на практике.

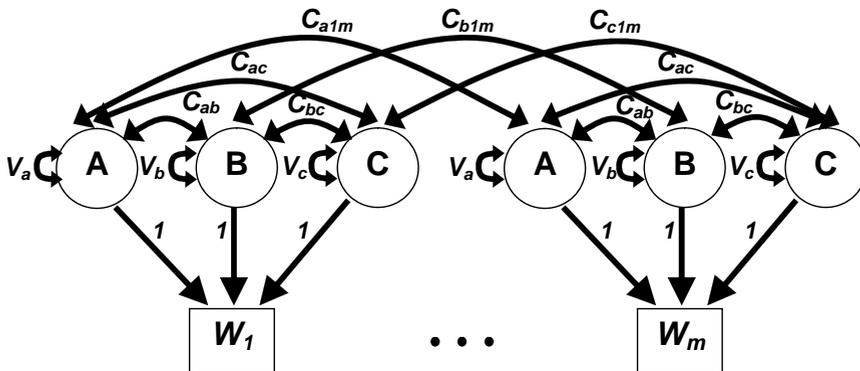


Рис. 4. Модель дисперсионных составляющих, представленная путевой диаграммой

⁶ Гипотеза о пропорциональности.

Интерпретация результатов рассмотренного варианта факторного анализа обычно опирается:

- на оценки свободных⁷ факторных дисперсий и ковариаций;
- оценки свободных корреляций между различными факторами в одни и те же моменты времени;
- оценки свободных корреляций между одинаковыми факторами в разные моменты времени;
- оценки статистической значимости различных компонентов модели.

2.4. Факторные модели и результаты расчетов

Данные для анализа содержали значения индекса коррупции K , индекса Джини D и показателя убийств U , измеренные в 22 европейских странах с 1998 по 2001 г.

Гипотеза о наличии общего фактора, влияющего на социально-экономические показатели, в первом приближении частично подтверждается оценками индекса надёжности Кронбаха, приведенными в табл. 1⁸.

Для последующего анализа, проведенного с помощью представленной на рис. 5 факторной модели дисперсионных составляющих, использовались вейвлет-коэффициенты, соответствующие периодам наблюдений с 1998 по 1999 г. и с 2000 по 2001 г. Вейвлет-коэффициенты, соответствующие периоду с 1998 по 1999 г., обозначены как KW_3 , DW_3 и UW_3 , а коэффициенты, соответствующие периоду с 2000 по 2001 год, – как KW_4 , DW_4 и UW_4 . Указан-

ная модель использовалась для выявления степени факторного взаимодействия и проверки гипотезы о наличии общего фактора, влияющего на наблюдаемые переменные.

В этой модели отсутствуют корреляции между факторами E_{D3} и E_{U3} , E_{D4} и E_{U4} , и E_{U3} и E_{D4} , поскольку их наличие приводит к неоднозначности решения, которая обусловлена появлением зависимостей между свободными параметрами модели и проявляется в дефекте матрицы системы уравнений, определяемой факторной моделью. Способ устранения зависимых переменных описан в статье об используемых методах альтернативного конфирматорного анализа [3].

Идентифицированные дисперсии и корреляции представлены на рис. 6⁹.

Для проверки гипотезы о наличии общего фактора F исследовалась значимость самого этого фактора, а также значимость ковариации данного фактора между первым и вторым периодом. В табл. 2 приведены значения статистики χ^2 для соответствующих упрощений модели и оценки статистической значимости исключенных компонентов.

Статистические значимости фактора F и его ковариации в разные периоды времени свидетельствуют о наличии общего фактора, влияющего на наблюдаемые переменные. Умеренная корреляция между значениями этого фактора в разные периоды наблюдений позволяет говорить о его временной изменчивости, обусловленной появлением новых составляющих.

Таблица 1

Группы переменных с высоким значением индекса надёжности Кронбаха

Группа переменных	Индекс надёжности Кронбаха
$U_{1998}, D_{1998}, K_{2000}$	0,52
$K_{1998}, U_{1999}, D_{2000}, D_{2001}$	0,63
$K_{1998}, U_{1999}, D_{2000}$	0,75

⁷ То есть представленных в модели в качестве свободных параметров.

⁸ Индексы переменных K , D и U обозначают год.

⁹ Идентифицированные ковариации между факторами для удобства интерпретации пересчитаны в корреляции.

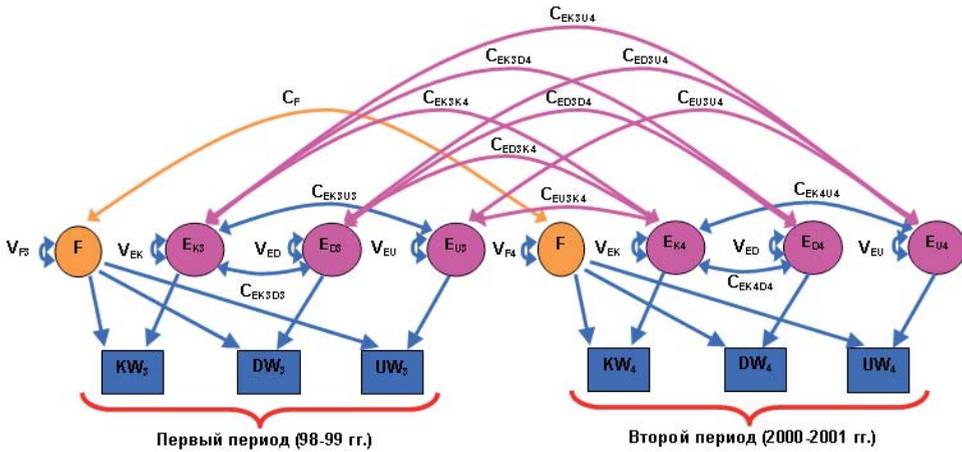


Рис. 5. Модель для проверки гипотезы о наличии общего фактора, влияющего на социально-экономические показатели в европейских странах: общий фактор обозначен как F; специфические факторы, влияющие на переменные KW., DW. и UW., – как E_K , E_D и E_U ; V_F – дисперсии общего фактора, V_{E_i} – дисперсии специфических факторов; $C_{E_i E_j}$ – ковариации между специфическими факторами; C_F – ковариация фактора F между первым и вторым периодом времени

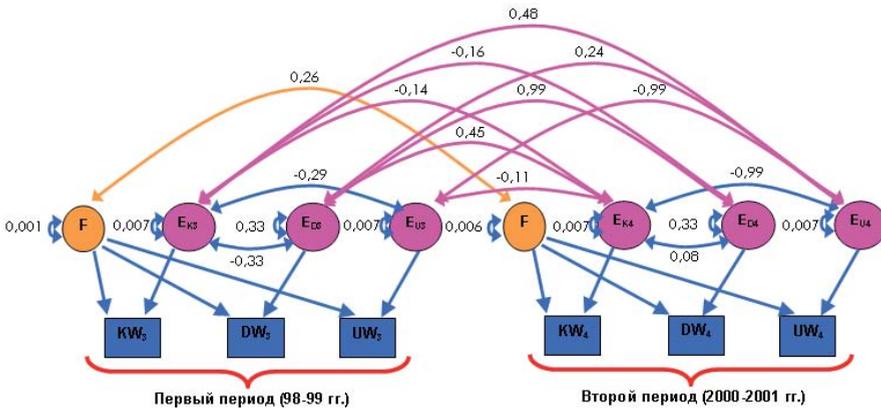


Рис. 6. Идентифицированные параметры факторной модели

Дальнейшие расчеты (табл. 3) показали незначительность влияния общего фактора на изменчивость индекса Джини. Однако удаление соответствующих влияний приводит к вырожденности матрицы переопределенной системы уравнений и появлению зависимых переменных, устранение которых вызывает качественное изменение типа системы и несогласованность полной и упрощенных моделей из табл. 2. Это позволяет говорить о целесообразности сохранения указанных связей.

2.5. Анализ с помощью факторных моделей: основные выводы

Факторный анализ индекса коррупции, индекса Джини и показателя убийств в европейских странах свидетельствует о наличии статистически значимого общего фактора, действием которого обусловлена изменчивость указанных показателей, причем влияние данного фактора подтверждено определенным временным изменением.

Таблица 2

Оценки статистической значимости исследуемых компонентов модели

Модель	Статистика χ^2	Число степеней свободы	p -значение	Разность оценок χ^2 по сравнению с полной моделью	Разность в числе степеней свободы по сравнению с полной моделью	p -значение для разности оценок χ^2	Значимость исключенного компонента
Модель с общим фактором F (полная модель)	6,764	3	0,054	–	–	–	–
Модель без ковариации между фактором F в 1-м и 2-м периодах наблюдений	11,195	4	0,024	4,430	1	0,035	Значим
Модель без общего фактора F	18,993	6	0,004	12,228	3	0,007	Значим

Таблица 3

Оценки балансов факторных влияний на изменчивость наблюдаемых переменных¹⁰

Баланс факторных влияний в первом периоде		Баланс факторных влияний во втором периоде	
Составляющие дисперсии KW_3		Составляющие дисперсии KW_4	
V_{F3}	17 %	V_{F4}	44 %
V_{EK}	87 %	V_{EK}	52 %
Необъясненная	0 %	Необъясненная	4 %
Составляющие дисперсии DW_3		Составляющие дисперсии DW_4	
V_{F3}	0 %	V_{F4}	3 %
V_{ED}	36 %	V_{ED}	163 %
Необъясненная	64 %	Необъясненная	0 %
Составляющие дисперсии UW_3		Составляющие дисперсии UW_4	
V_{F3}	27 %	V_{F4}	24 %
V_{EU}	132 %	V_{EU}	26 %
Необъясненная	0 %	Необъясненная	50 %

¹⁰ Дисперсии общего и специфического факторов указаны в процентах от наблюдаемой дисперсии соответствующей переменной. Из-за наличия невязок сумма дисперсий общего и специфического факторов может не совпадать с наблюдаемой дисперсией.

Используя результаты, представленные в табл. 3 и на рис. 6, можно сделать следующие выводы о балансе влияний общего и специфических факторов на каждый показатель, а также об изменении факторных влияний во времени:

- индекс коррупции и показатель убийств испытывают сопоставимые влияния общего и специфического факторов, причем в 1998–99 гг. влияние специфических факторов на изменчивость указанных характеристик значительно преобладает, а в 2000–01 гг. указанные влияния выравниваются;

- в случае индекса Джини влияние общего фактора пренебрежимо мало по сравнению с влиянием специфического фактора, причем в 2000–01 гг. специфический фактор практически полностью определяет всю изменчивость данной характеристики;

- имеет место существенная отрицательная корреляция между специфическими факторами, воздействующими на индекс коррупции и показатель убийств, усилившаяся во втором периоде наблюдений;

- имеет место сильная отрицательная корреляция между специфическими факторами, воздействующими на показатель убийств в первом и втором периоде наблюдений;

- имеет место сильная положительная корреляция между специфическими факторами, воздействующими на индекс Джини в первом и втором периоде наблюдений.

3. Сети Маркова

Второй метод анализа – «Сети Маркова» – рассматривается здесь как дополнение к представленному выше подходу [11]. Его применение позволяет исследовать динамику развития и взаимосвязь исследуемых характеристик, восстанавливать вероятностные характеристики исследуемых величин в произвольные моменты времени, находящиеся между контрольными точками, а также извлекать из экспериментальных результатов новые показатели, недоступные для других методов исследования. Непосредственно анализируемыми

объектами являются гистограммы, представляющие распределения наблюдаемых значений исследуемых характеристик в контрольные моменты времени.

3.1. Основные принципы построения

Анализируемой величиной являются диапазоны значений (уровни) исследуемой характеристики. Весь диапазон допустимых значений заданной величины делится на несколько интервалов. Каждый интервал рассматривается как отдельное состояние, в котором исследуемая величина может находиться с некоторой вероятностью, отражающей частоту ее появления в популяции, переходя из одного состояния в другое по определенным правилам. Для описания, как вероятности нахождения в заданных состояниях изменяются со временем, применяются сети Маркова.

По результатам накопленных наблюдений проводится прогнозирование указанных вероятностей с использованием параметрических математических моделей, описываемых *марковскими случайными процессами с дискретными состояниями и непрерывным временем*.

Модель для описания динамики этих переходов представляется ориентированным графом, в котором вершины соответствуют состояниям, а дуги соответствуют переходам. Процесс переходов между состояниями может рассматриваться как случайное блуждание по графу из одного состояния в другое согласно направлениям дуг. Переходы между состояниями мгновенны и происходят в случайные моменты времени.

Можно показать, что в рассматриваемых потоках число событий X , попадающих в любой временной интервал длины τ , начинающийся в момент t , распределено согласно *закону Пуассона*:

$$P_{t,\tau}(X=m) = \frac{(a(t,\tau))^m}{m!} e^{-a(t,\tau)},$$

где $P_{t,\tau}(X=m)$ – вероятность появления m событий в течение рассматриваемого интервала, $a(t, \tau)$ – среднее число событий, попадающих в интервал длины τ , начинающийся в момент времени t .

Далее будут рассматриваться только *стационарные потоки* (в которых $\mathbf{a}(t, \tau) = \eta \tau$, $\eta = \text{const}$). Параметр η называется *интенсивностью стационарного потока*. Он равен среднему числу событий в единицу времени. Средняя продолжительность времени между двумя смежными событиями в этом случае равна $1/\eta$.

При указанных предположениях динамика изменения во времени вероятностей пребывания системы в различных состояниях описывается системой обыкновенных дифференциальных уравнений Колмогорова, в которой каждому состоянию соответствует уравнение

$$\frac{dp_k(t)}{dt} - \sum_j \eta_{kj} p_j(t) + \sum_i \eta_{ik} p_i(t),$$

где k – номер состояния; P_k и P_i – вероятности нахождения в k -м и i -м состояниях; η_{kj} – интенсивности потоков, выходящих из состояния k ; η_{ik} – интенсивности потоков, входящих в состояние k .

Для интегрирования указанной системы уравнений необходимо задать начальные условия

$$p_0(0), p_1(0), \dots, p_n(0); \sum_{k=0}^n p_k(0) = 1.$$

Нормализующее уравнение $\sum_{k=0}^n p_k(t) = 1$ выполняется в любой момент времени.

Интенсивности потоков событий являются свободными параметрами модели. Здесь рассматривается обратная постановка задачи. Значения свободных параметров определяются путем сравнения наблюдаемых и прогнозируемых гистограмм, описывающих наблюдаемые распределения исследуемой характеристики, а именно: вычисляются значения, обеспечивающие наилучшее соответствие наблюдаемых и ожидаемых частот попадания в определенное состояние системы в заданные моменты времени. Прогнозируемые вероятности нахождения в состояниях получаются путем численного интегрирования систем уравнений Колмогорова.

Прогнозируемая частота попадания в k -е состояние равна $p_k N$, где N – число

элементов в выборке. Соответствующие наблюдаемые частоты F_k определяются по распределениям наработок до отказа, зарегистрированным на практике. *Статистика Пирсона*

$$\chi^2 = \sum_{k=0}^n \frac{(F_k - p_k N)^2}{p_k N}$$

используется как мера соответствия в том смысле, что ее большие значения означают плохое согласование прогнозируемых и наблюдаемых результатов, а малые значения – хорошее согласование. Сумма этих статистик минимизируется в те моменты времени, для которых имеются результаты наблюдений.

Данный способ идентификации свободных параметров называется *методом минимума χ^2* [1] и дает решения, близкие к полученным ММП [1, с. 461–462]. Это позволяет использовать его для проверки гипотезы, что полученный прогноз согласуется с результатами наблюдений.

3.2. Многофакторные сети Маркова

Для того чтобы модель более детально представляла исследуемые характеристики, удобно использовать двух- и многофакторные сети Маркова [9]. Типовая структура двухфакторной сети показана на рис. 7. Данная модель учитывает два определяемых исследователем фактора различной природы, значимых для исследуемой характеристики. Каждый фактор соответствует одному из двух ортогональных направлений на плоскости: вертикальному или горизонтальному. Каждая строка (столбец) состояний сети соответствует определенному уровню значений фактора. Переходы между состояниями возможны только вдоль указанных ортогональных направлений («манхэттенские переходы»), а именно: интенсивности λ_{ij} соответствуют правым горизонтальным переходам, интенсивности μ_{ij} – левым, интенсивности θ_{ij} соответствуют вертикальным переходам вниз, а интенсивности η_{ij} – вертикальным переходам вверх. В общем случае эти интенсивности переходов являются функциями их позиции в диаграмме

$$\lambda_{ij} = \lambda(i, j), \mu_{ij} = \mu(i, j), \theta_{ij} = \theta(i, j), \eta_{ij} = \eta(i, j).$$

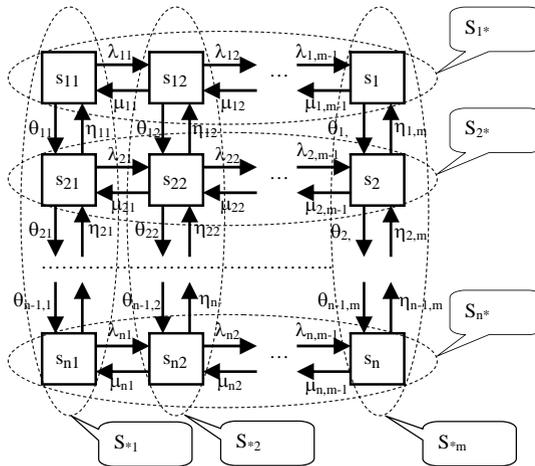


Рис. 7. Двухфакторная сеть Маркова

Подобные функции позволяют отразить в модели тонкие особенности поведения исследуемой системы.

В практически важных случаях идентификация свободных параметров двух- и многофакторных моделей является сложной вычислительной задачей. Однако имеется эффективный прием, позволяющий упростить эту работу. Рассмотрим его на примере двухфакторной модели.

Строки и столбцы состояний исходной сети могут быть объединены в «сгруппированные» состояния, а именно: нахождение в состоянии S_i^* ($i=1, 2, \dots, n$) равносильно нахождению в одном из состояний $s_{i1}, s_{i2}, \dots, s_{im}$, а нахождение в состоянии S_j^* ($j=1, 2, \dots, m$) равносильно нахождению в одном из состояний $s_{1j}, s_{2j}, \dots, s_{nj}$. Согласно теоремам о суммах пуассоновских потоков событий, переходы между сгруппированными состояниями также удовлетворяют свойствам пуассоновских потоков событий, а их интенсивности являются суммами аналогичных характеристик в соответствующих направлениях:

$$\Lambda_i = \sum_{k=1}^n \lambda_{ki}, \quad M_i = \sum_{k=1}^n \mu_{ki}, \quad \Theta_i = \sum_{k=1}^m \theta_{ik}, \quad H_i = \sum_{k=1}^m \eta_{ik}.$$

Таким образом, исследование исходной двухфакторной системы сводится к анализу двух производных однофакторных систем (рис. 8 и 9), которые могут быть идентифицированы и интерпретированы, как показано выше. Возврат к исходной систе-

ме прост: вероятность p_{ij} нахождения в состоянии s_{ij} может быть вычислена как произведение вероятностей P_i и P_j пребывания в «сгруппированных» состояниях S_i^* и S_j^* , соответственно. Наличие такого простого способа декомпозиции является важным преимуществом рассматриваемых систем. В практических приложениях представленная выше общая модель может быть упрощена и конкретизирована.

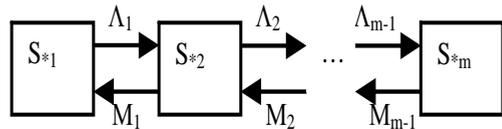


Рис. 8. Производная однофакторная система (результат группировки столбцов состояний)

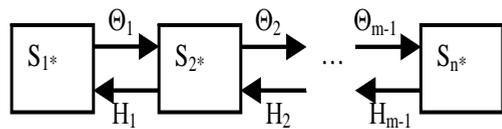


Рис. 9. Производная однофакторная система (результат группировки строк состояний)

3.3. Исследование с помощью двухфакторных сетей Маркова

Для исследования взаимодействия и эволюции факторов, определяющих ИНСО, использованы двухфакторные сети Маркова. Исследовался период времени с 2000

по 2006 г. Построены три двухфакторных модели, выявляющие попарное взаимодействие и развитие трех составляющих, определяющих индекс нравственности (рис. 10):

- модель индекс Джини/уровень коррупции;
- модель индекс Джини/уровень убийств;
- модель уровень коррупции/уровень убийств.

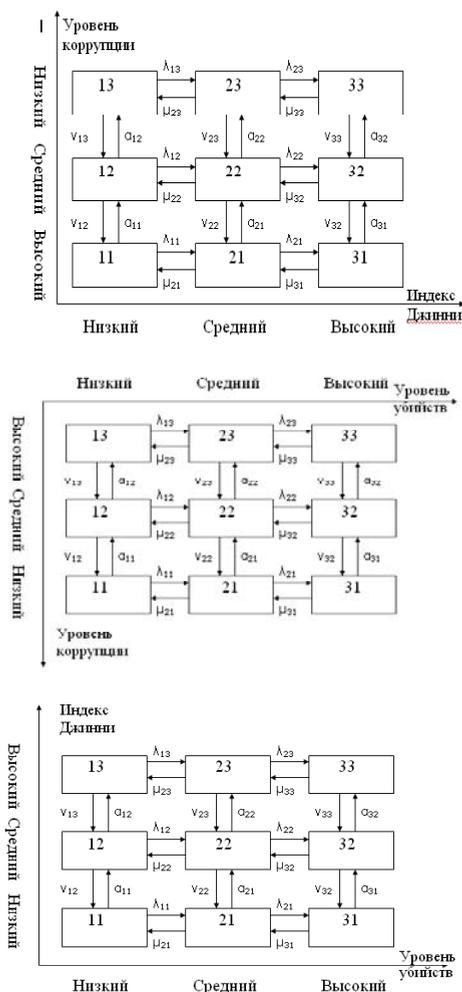


Рис. 10. Двухфакторные модели для исследования взаимодействий рассматриваемых характеристик

Особенности моделей:

- каждая страна должна быть оценена по двум факторам, включенным в модель;
- каждое ортогональное направление включает в себя три состояния, так как именно это количество позволяют исходные данные;
- начальными распределениями моделей являются наблюдаемые распределения в предыдущей контрольной точке (в 2000 г.);
- неявный параметр системы – время. Время измеряется в годах;
- свободными параметрами системы являются параметры показателей интенсивности перехода.

Взаимные связи интенсивностей потоков событий описываются следующими формулами:

$$\left\{ \begin{aligned} \lambda_{ij} &= \lambda_{i_0 j_0} I^{\Delta i} Y^{\Delta j}; \\ \mu_{ij} &= \mu_{i_0 j_0} I^{\Delta i} Y^{\Delta j}; \\ \alpha_{ij} &= \alpha_{i_0 j_0} I^{\Delta i} Y^{\Delta j}; \\ v_{ij} &= v_{i_0 j_0} I^{\Delta i} Y^{\Delta j}, \quad \Delta i = i - i_0, \quad \Delta j = j - j_0. \end{aligned} \right.$$

$\lambda_{i_0 j_0}, \mu_{i_0 j_0}, \alpha_{i_0 j_0}, v_{i_0 j_0}$ – «начальные» интенсивности. Начальной полагается нижняя левая интенсивность. Например, $v_{i_0 j_0} = v_{12}$, где i – номер состояния по горизонтали, а j – по вертикали.

Таким образом, свободными идентифицируемыми параметрами являются «начальные» интенсивности и

$$I_\lambda, Y_\lambda, I_\mu, Y_\mu, I_\alpha, Y_\alpha, I_v, Y_v.$$

Если

$I > 1$, то с ростом i увеличится интенсивность потоков вниз (v);

$I < 1$, то с ростом i увеличится интенсивность потоков вверх (α);

$Y > 1$, то с ростом j уменьшится интенсивность потоков вниз (v);

$Y < 1$, то с ростом j уменьшится интенсивность потоков вверх (α).

При девяти состояниях в разработанной системе соотношений минимизировалась

сумма двух статистик Пирсона, вычисленных в двух контрольных точках. Проведенные расчеты показали хорошее согласование прогнозируемых и наблюдаемых гистограмм.

В модели индекс Джини/уровень коррупции значение суммарной статистики Пирсона равно 1,564 ($df=5, p=0,906$). Для модели индекс Джини/уровень убийств этот показатель составляет 3,064 ($df=5, p=0,69$), а для модели уровень убийств/уровень коррупции – 2,115 ($df=5, p=0,833$). Полученные характеристики свидетельствуют о достаточно высокой степени соответствия модели наблюдениям.

3.4. Анализ с помощью сетей Маркова: основные выводы

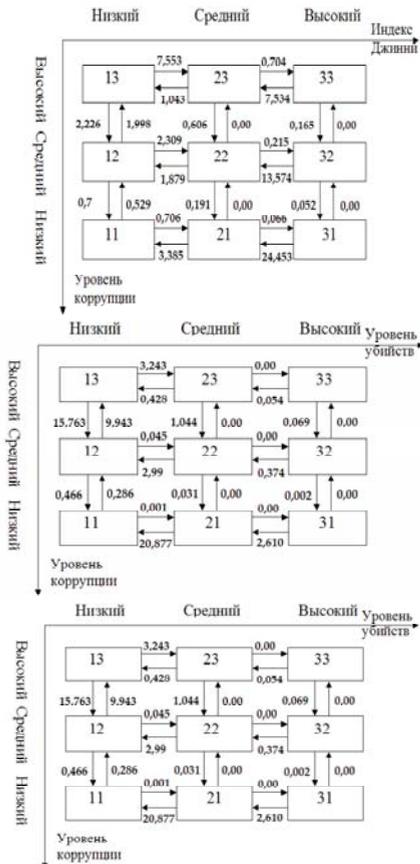


Рис. 11. Интенсивности потоков событий в двухфакторных моделях

Полученные оценки интенсивностей потоков событий (рис. 11) позволяют сделать следующие выводы.

1. Модель «уровень коррупции/уровень убийств». Тенденция к уменьшению уровня коррупции преобладает над тенденцией к его увеличению. При этом чем ниже уровень убийств, тем быстрее будет уменьшаться коррупция. При низких и средних значениях показателя уровня коррупции существует тенденция к уменьшению числа убийств, в отличие от высокого уровня коррупции, при котором уровень убийств стремится к средним значениям.

2. Модель «уровень коррупции/индекс Джини». Значения индекса Джини стремятся к средним показателям при средних и высоких значениях уровня коррупции. При низких показателях уровня коррупции видна тенденция к уменьшению индекса Джини. Тенденция к уменьшению уровня коррупции преобладает над тенденцией к его увеличению, при этом чем меньше индекс Джини, тем быстрее уменьшается коррупция.

3. Модель «индекс Джини/уровень убийств». Тенденция к уменьшению числа убийств преобладает над тенденцией к его увеличению. Значения индекса Джини имеют тенденцию к уменьшению при низких показателях уровня убийств. При высоком и среднем уровне убийств существует тенденция к увеличению значений индекса Джини, при этом чем больше был индекс Джини, тем быстрее.

Таким образом, осуществленный анализ продемонстрировал правомерность использования разработанного Институтом психологии РАН индекса нравственного состояния общества, а также выветил взаимосвязи его составляющих.

Литература

1. Крамер Г. Математические методы статистики. М., 1976.
2. Куравский Л. С., Корниенко П. А. Применение нейронных сетей для идентификации локусов количественных признаков в психогенетике // Нейрокомпьютеры: разработка и применение. 2007. № 4.
3. Куравский Л. С., Мармалюк П. А., Баранов С. Н., Абрамочкина В. И., Петрова Е. А. Факторный анализ результатов вейвлет-преобразований лонгитюдных данных как новый метод исследования динамических характеристик сложных систем // Нейрокомпьютеры: разработка и применение. 2009. № 9.
4. Юревич А. В. Нравственность как психологическая проблема // Вопросы психологии. 2009. № 4.
5. Юревич А. В., Ушаков Д. В. Макропсихология как новая область психологических исследований // Вопросы психологии. 2007. № 4.
6. Юревич А. В., Ушаков Д. В., Цапенко И. П. Количественная оценка макропсихологического состояния современного российского общества // Психологический журнал. 2007. № 4.
7. Bollen K. A. Structural equations with latent variables. N. Y., 1989.
8. Jöreskog K. G. Estimation and testing of simplex models // British Journal of Mathematical and Statistical Psychology. Vol. 23.
9. Kuravsky L. S., Baranov S. N. The concept of multifactor Markov networks and its application to forecasting and diagnostics of technical systems // In: Proc. Condition Monitoring 2005, Cambridge, United Kingdom, 2005.
10. Kuravsky L. S. and Baranov S. N. Development of the wavelet-based confirmatory factor analysis for monitoring of system factors // In: Proc. 5th International Conference on Condition Monitoring & Machinery Failure Prevention Technologies, Edinburgh, United Kingdom, 2008.
11. Kuravsky L. S. and Baranov S. N. Neural networks in fatigue damage recognition: diagnostics and statistical analysis // In: Proc. 11th International Congress on Sound and Vibration, St.-Petersburg, Russia, 2004.
12. Kuravsky L. S., Kornienko P. A. On the approach to identifying quantitative trait loci in behavior genetics. In: Proc. 2nd World Congress on Engineering Asset Management and 4th International Conference on Condition Monitoring, Harrogate, United Kingdom, 2007.
13. Neale M. C., Cardon L. R. Methodology for genetic studies of twins and families. Dordrecht, the Netherlands, Kluwer Academic Publishers, 1992.

Factor Analysis of Society's Moral State Indices in European Countries

L. S. Kuravsky,

PhD in Engineering, Professor, Head of the Applied Informatics Chair; Dean, Department of Information Technologies, Moscow State University of Psychology and Education

A. V. Yurevich,

PhD in Psychology, Professor, Corresponding Member of the Russian Academy of Sciences, Vice Director of the Psychological Institute of Russian Academy of Sciences

P. A. Marmalyuk,

Software Engineer, Applied Informatics Chair, Department of Information Technologies, Moscow State University of Psychology and Education

E. G. Ivanova,

Psychologist, Applied Informatics Chair, Department of Information Technologies, Moscow State University of Psychology and Education

The paper presents the instrument and the results of factor analysis of developmental and interaction characteristics dynamics that define the index of the moral state of society, developed by the Institute of Psychology, Russian Academy of Sciences. The alternative way of confirmatory factor analysis and Trained Markov Networks revealed a number of developmental patterns of problematic characteristics that indicate usage appropriateness of this index. The advantages of the applied type of factor analysis include: the possibility of finding estimates of free parameters of the model via direct methods (and not by iteration), that guarantee an unambiguous and optimal solution, flexible means of studying the factor interactions and the applicability of covariance and correlation matrices of arbitrary structure for the analysis. Markov Networks, in turn, revealed new types of indicators that are inaccessible to conventional approaches.

Keywords: morality, the index of the moral state of society, factor analysis, the Markov Networks.

Литература

1. *Kramer G.* Matematicheskie metody statistiki. M., 1976.
2. *Kuravskii L. S., Kornienko P. A.* Primenenie neironnyh setei dlya identifikatsii lokusov kolichestvennykh priznakov v psihogenetike. Neurokomp'yutery: razrabotka i primeneniye. 2007. № 4.
3. *Kuravskii L. S., Marmalyuk P. A., Baranov S. N., Abramochkina V. I., Petrova E. A.* Faktorniy analiz rezul'tatov veivlet-preobrazovaniya longitudnykh dannykh kak novyi metod issledovaniya dinamicheskikh harakteristik slozhnykh sistem. Neurokomp'yutery: razrabotka i primeneniye. 2009. № 9.
4. *Yurevich A. V.* Nравstvennost' kak psihologicheskaya problema // Voprosy psihologii. 2009. № 4.
5. *Yurevich A. V., Ushakov D. V.* Makropsihologiya kak novaya oblast' psihologicheskikh issledovaniy // Voprosy psihologii. 2007. № 4.
6. *Yurevich A. V., Ushakov D. V., Capenko I. P.* Kolichestvennaya ocenka makropsihologicheskogo sostoyaniya sovremennogo rossiyskogo obshchestva // Psihologicheskii zhurnal. 2007. № 4.
7. *Bollen K. A.* Structural equations with latent variables. New York, 1989.
8. *Jöreskog K. G.* Estimation and testing of simplex models. British Journal of Mathematical and Statistical Psychology. Vol. 23.
9. *Kuravsky L. S., Baranov S. N.* The concept of multifactor Markov networks and its application to forecasting and diagnostics of technical systems // In: Proc. Condition Monitoring 2005, Cambridge, United Kingdom, 2005.
10. *Kuravsky L. S. and Baranov S. N.* Development of the wavelet-based confirmatory factor analysis for monitoring of system factors // In: Proc. 5th International Conference on Condition Monitoring & Machinery Failure Prevention Technologies, Edinburgh, United Kingdom, 2008.
11. *Kuravsky L. S. and Baranov S. N.* Neural networks in fatigue damage recognition: diagnostics and statistical analysis // In: Proc. 11th International Congress on Sound and Vibration, St.-Petersburg, Russia, 2004.
12. *Kuravsky L. S., Kornienko P. A.* On the approach to identifying quantitative trait loci in behavior genetics. In: Proc. 2nd World Congress on Engineering Asset Management and 4th International Conference on Condition Monitoring, Harrogate, United Kingdom, 2007.
13. *Neale M. C., Cardon L. R.* Methodology for genetic studies of twins and families. Dordrecht, the Netherlands, Kluwer Academic Publishers, 1992.