

ПРИМЕНЕНИЕ МЕТОДОВ СИСТЕМНОГО АНАЛИЗА ДЛЯ УСТАНОВЛЕНИЯ ВЛИЯНИЯ СКРЫТЫХ ФАКТОРОВ НА СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ ПОКАЗАТЕЛИ ГОСУДАРСТВЕННОГО РАЗВИТИЯ

УДК 004.921

КОЛЯДА Юрий Евгеньевич

д.ф.-м.н., проф., заведующий кафедрой математических методов и системного анализа,
Мариупольский государственный университет.

Научные интересы: системный анализ, исследование систем, проектирование компьютерных систем поддержки принятия решений, обработка изображений.

МЕРКУЛОВА Екатерина Владимировна

к.т.н., доцент, доцент кафедры математических методов и системного анализа,
Мариупольский государственный университет.

Научные интересы: системный анализ, исследование систем, проектирование компьютерных систем поддержки принятия решений, обработка изображений.

ВВЕДЕНИЕ

Возникновение государства, его развитие, расцвет и упадок, а в некоторых случаях и исчезновение с карты мира, обусловлено, как известно, экономическими, политическими и социальными процессами [1-3].

Анализ государственного развития последних тысячелетий констатирует стремительное развитие некоторых государств, а затем потерю ими лидирующих позиций на мировой арене. За последние сотни лет в связи с освоением новых континентов и политико-экономических преобразований появились новые государства. В некоторых из них в настоящее время достигнут высокий уровень экономического развития, а в других рост экономического благосостояния происходит достаточно медленно, хотя стартовые условия и природные ресурсы были у всех, практически, одинаковы. Многие государства с тысячелетней историей также имеют различный уровень экономического развития.

Почему так происходит? Почему так различен жизненный уровень большинства государств? В этой связи возникает вопрос: «А существуют ли более глубокие причины, чем упомянутые выше, обуславливающие

закономерности государственного становления и развития?» Резонно сделать предположение об их существовании. Они могут быть скрытыми, недоступными для анализа и измерений и, вполне возможно, влияющими на экономические, политические и социальные процессы в обществе. На возможность их существования с свое время указывал Кант [4]. В частности, он предполагал, что «...идиотическое течение всего человеческого не имеет никакой видимой закономерности и, что человеческая история кажется непрерывной цепью войн и жестокостей». Но всё же он ставил вопрос: «Нет ли в истории человечества некоего регулярного движения, такого, которое кажется хаотическим с точки зрения индивида, но в котором можно обнаружить медленную и прогрессивную долговременную тенденцию?».

Вполне возможно, что именно благодаря наличию скрытых факторов может устанавливаться причинно-следственная связь между, кажущимися на первый взгляд хаотическими, наблюдаемыми количественными признаками, характеризующими государство. Если государство рассматривать как систему, то для анализа эффективности функционирования такой системы,

можно применить методы системного анализа и, в качестве указанных количественных признаков, или величин, предлагается использовать геополитические характеристики, поддающиеся непосредственному измерению, а также валовой внутренний продукт (ВВП) на душу населения. Последний показатель, собственно, и определяет экономический уровень развития государства.

Поэтому целью настоящей работы является применение методов системного анализа для оценки эффективности функционирования государства как системы и установления связей между элементами такой системы, а именно геополитическими характеристиками и экономическим уровнем развития государства, как системы, в предположении существования скрытых, недоступных для прямых измерений, факторов.

ПОСТАНОВКА ЗАДАЧИ

Для исследования функционирования нашей системы и изучения вопроса выявления связей между геополитическими характеристиками и экономическим уровнем развития государства предлагается применить методы системного анализа [5], а именно комбинированный математический аппарат факторного и статистического анализа [6,7]. Суть его заключается в многомерном статистическом анализе, казалось бы, хаотических, наблюдаемых величин и формировании на их основе корреляционной матрицы. Предполагается, что имеющая место множественная корреляция между наблюдаемыми величинами обусловлена не только их взаимными связями, но и посредством невидимых, гипотетических факторов, не поддающихся непосредственным наблюдениям и измерениям [8, 9]. При этом необходимо установить и выделить указанные факторы по данным статистических наблюдений, так, чтобы с их помощью оказалось возможным воспроизвести и достаточно точно описать наблюдаемые величины и зависимости между ними. Количество факторов может быть значительно меньше количества наблюдаемых величин [10]. Это позволит упорядочить кажущуюся хаотичность социальных процессов путём сжатия информации и поможет создать более адекватную мо-

дель явлений и связей, происходящих в государстве и обществе и глубже понять их суть.

В качестве наблюдаемых и поддающихся непосредственным измерениям количественных признаков государств использовались следующие геополитические характеристики: возраст государства, его площадь, количество населения, средняя годовая температура и ВВП.

РЕШЕНИЕ ЗАДАЧИ

Для анализа были выбраны 50 стран со всех континентов. В таблице 1 приведены эти государства и их геополитические характеристики, а также ВВП. Следует отметить, что при составлении указанной таблицы авторы испытывали определенные трудности, так как в разных источниках значения для одного и того же показателя несколько отличались.

Приведенные данные представляют собой кажущиеся хаотичными наблюдаемые величины, или переменные.

Алгоритм обработки переменных осуществляется в соответствии с методикой факторного анализа, описанной в [4, 5, 8]. В частности, исходные данные были представлены в виде матрицы переменных $Y=(y_{ij})$, где $i=1,2,\dots,m$ относится к переменным, а $j=1,2,\dots,n$ – к странам. В нашем случае $n=50$, а $m=5$. Коэффициент корреляции между двумя переменными i и k вычислялся по известной формуле:

$$\begin{aligned} n_{ik} &= \frac{\sum_{j=1}^n (y_{ij} - \bar{y}_i) (y_{kj} - \bar{y}_k)}{\sqrt{\sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 \sum_{j=1}^n (y_{kj} - \bar{y}_k)^2}} \\ &= \frac{s_{ik}}{s_i s_k} \end{aligned} \quad (1)$$

где $\bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}$ – среднее значение;

$s_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}$ – стандартное

отклонение, аналогично и для s_k ;

Табл.1.

Исходные данные

№ п/п	Названия стран	Возраст государства (Vo _{zr})	Площадь (P _{los})	Количество населения (N _{as})	ВВП (VVP)	Средняя годовая температура (Temp)
1	Китай	3500	9598077	1322178190	7222	26
2	Индия	1409	3287590	1131191071	3802	27
3	США	232	9532776	303097000	43223	11.2
4	Бразилия	509	8511965	190256710	10073	23.5
5	Нигерия	537	923768	146300000	1227	25
6	Япония	2668	377835	127417000	32530	14.8
7	Мексика	198	1972550	105700891	11369	23
8	Вьетнам	1909	329560	85395893	3393	27
9	Египет	5009	1001450	78329784	4895	27
10	Таиланд	770	514000	65095420	9193	32
11	Южная Корея	64	98480	49034812	24084	10
12	ЮАР	357	1219912	44074732	13018	25.1
13	Марокко	909	446550	33785853	5765	18.5
14	Алжир	2209	2381740	33322179	7747	18
15	Канада	409	2381740	33004202	35514	5.1
16	Румыния	1496	237500	22284801	10125	9.8
17	Казахстан	2509	2727300	15401170	9568	10
18	Тунис	304	163610	10299331	8975	17.5
19	Швеция	709	449964	9031088	34735	6.25
20	Швейцария	360	41290	7519318	38706	10.75
21	Болгария	1328	110910	7389678	10022	13
22	Израиль	61	26990	7300474	31561	20
23	Ливия	57	1759540	6046124	12848	25
24	Дания	1208	43094	5442996	36920	16
25	Финляндия	428	337030	5238460	35559	5.2
26	Украина	18	603700	46156985	3995	19
27	Россия	900	17075400	141888900	9074	5.9
28	Беларусь	18	207600	9675800	10900	6
29	Литва	18	65200	3366200	14273	10.95
30	Греция	180	131940	11300000	20081	25
31	Германия	1509	357022	8230000	40415	11
32	Великобритания(Англия)	943	133395	60776238	46432	17
33	Франция	1165	547030	6190000	45858	16
34	Португалия	609	92391	10642836	18398	22
35	Испания	516	504782	40847371	35557	22
36	Италия	2000	301230	60000000	39565	15
37	Польша	1049	312679	38636157	11860	7.5
38	Венгрия	1009	93030	10059000	11218	11
39	Венесуэла	510	916445	26400000	11933	27
40	Никарагуа	507	129494	5800000	3100	25
41	Аргентина	409	2766890	40301927	8146	22
42	Чили	457	756950	16000000	14673	13
43	Австралия	403	7686850	20001546	37000	25
44	Новая Зеландия	169	268680	4284000	34121	13
45	Грузия	1710	69700	4630003	3119	15
46	Иран	2309	1648000	71208000	12300	15
47	ОАЭ	1409	83600	4500000	21100	32
48	Турция	2809	780580	71158647	10737	24
49	Норвегия	1309	385199	4799252	72305.6	5.2
50	Исландия	1028	103125	319756	40277	4

$s_{ik} = \frac{1}{n-1} \sum_{j=1}^n (y_{ij} - \bar{y}_i)(y_{kj} - \bar{y}_k)$ – ковариация. Затем элементы матрицы Y преобразовывались к стандартизованным или нормированным переменным с соответствующим образованием матрицы $Z=(z_{ij})$, элементы которой имеют вид:

$$z_{ij} = \frac{y_{ij} - \bar{y}_i}{s_i} \quad (2)$$

При этом матрица $Z=(z_{ij})$ удовлетворяет следующим условиям:

$$r_{ik} = \frac{\sum_{j=1}^n (z_{ij}z_{kj})}{\sqrt{\sum_{j=1}^n z_{ij}^2 \sum_{j=1}^n z_{kj}^2}} = s_{ik} = \frac{1}{n-1} \sum_{j=1}^n z_{ij}z_{kj} \quad (5)$$

Из (5) следует, что для стандартизованных переменных коэффициенты корреляции и ковариации равны. Переход к стандартизованным переменным позволяет получить еще одно важное соотношение:

$$\frac{1}{n-1} ZZ' = R = S \quad (6)$$

Здесь R и S корреляционная и ковариационная матрицы соответственно, Z' транспонированная матрица стандартизованных переменных.

Целью любого метода факторного анализа является представление величин z_{ij} , т.е. элементов матрицы Z , в виде линейной комбинации нескольких гипотетических переменных, или факторов, число которых равно r .

Тогда справедливо соотношение:

$$z_{ij} = a_{i1}p_{1j} + a_{i2}p_{2j} + \dots + a_{ir}p_{rj} \quad (7)$$

Это равенство выражает основную модель факторного анализа. Здесь a_{il} – постоянные коэффициенты, которые следует определить; p_{lj} – значения факторов у j -го государства, которые также неизвестны.

Соотношение (7) в матричной форме может быть представлено в виде:

$$Z=AP \quad (8)$$

Следует напомнить, что Z является матрицей стандартизованных переменных – исходных данных,

$$\frac{1}{n} \sum_{j=1}^n z_{ij} = 0, \quad i=1,2,\dots,m \quad (3)$$

$$\frac{1}{n-1} \sum_{j=1}^n z_{ij}^2 = 1, \quad i=1,2,\dots,m \quad (4)$$

Т.е. все средние значения переменных матрицы Z равны нулю, а все дисперсии – единице. Тогда в стандартизованных переменных формула (1) упрощается:

порядка $m \times n$, $A=(a_{il})$ является неизвестной матрицей порядка $m \times r$. Она называется факторным отображением, а ее элементы (коэффициенты a_{il}) – факторными нагрузками. Т.е. A является матрицей коэффициентов регрессии факторов по переменным. $P=(p_{lj})$ – неизвестная матрица порядка $r \times n$ значений всех факторов у всех государств, или факторные значения.

Таким образом, в уравнении (8) известна лишь матрица Z , а матрицы A и P – неизвестны. Уравнение (8) без введения дополнительных ограничений имеет бесконечное множество решений. Эти ограничения для всех методов факторного анализа сводятся к следующим. Отдельные наблюдаемые значения являются линейной комбинацией гипотетических, ненаблюдаемых, или скрытых переменных, называемых факторами, которые не могут быть обнаружены в процессе наблюдения. Тогда подставив в равенство (8), являющееся основной моделью факторного анализа, уравнение (6), получим:

$$R = \frac{1}{n-1} ZZ' = \frac{1}{n-1} AP(AP)' = \frac{1}{n-1} APP'A' = A \frac{1}{n-1} PP'A' \quad (9)$$

Теперь по аналогии с формулой (6) можно утверждать, что выражение $\frac{1}{n-1}PP' = C = (c_{ip})$ является корреляционной матрицей, отражающей связи между факторами, т.е.:

$$R=ACA' \quad (10)$$

Если на это равенство наложить условие некоррелированности факторов, т.е. постулируются ортогональные факторы, то матрица $C=1$ становится единичной, и тогда (8) будет иметь вид:

$$R=AA' \quad (11)$$

Соотношения (10) и (11) называют фундаментальной теоремой факторного анализа. Она утверждает, что корреляционная матрица может быть воспроизведена с помощью факторного отображения и корреляций между факторами, если таковые имеются. В дальнейшем будем предполагать ортогональность факторов. Таким образом, соотношения (6), (8) и (11) позволяют найти неизвестные матрицы A и P , воспроизвести корреляционную матрицу путём сжатия информации, что значительно упростит системный анализ происходящих процессов и явлений.

Задача решалась численно. Ниже приведены результаты полученных вычислений.

В таблице 2 представлена корреляционная матрица исходных данных – наблюдаемых величин.

Табл.2.

Корреляционная матрица исходных данных

Переменная	Vozi	Plos	Nas	VVP	Temp
Vozi	1,00	0,03	0,32	-0,20	0,17
Plos	0,03	1,00	0,40	-0,00	-0,17
Nas	0,32	0,40	1,00	-0,20	0,26
VVP	-0,20	-0,00	-0,20	1,00	-0,35
Temp	0,17	-0,11	0,26	-0,35	1,00

Из нее следует, что между ними наблюдаются корреляционные связи от слабой – с коэффициентами корреляции $r_{ij} < |0,25|$ до умеренной (значимой) с $|0,25| < r_{ij} < |0,75|$. Данная корреляционная таблица достаточно информативна. Из нее видно, что имеются как положительные, так и отрицательные коэффициенты корреляции. Например, первый показатель, возраст государства, практически не коррелирует с его площадью, но имеет положительный коэффициент с численностью населения, коэффициент корреляции при этом немал, равен 0,32. Переменные площадь государства и численность населения коррелируют с небольшим коэффициентом, равным 0,4. Особо следует обратить внимание на такой важный показатель, как ВВП. Он имеет отрицательные коэффициенты корреляции со всеми переменными, равен нулю лишь с переменной площадь государства. Обращает на себя внимание коэффициент корреляции между ВВП и среднегодовой температурой. Он по абсолютной величине максимален в таблице, равен 0,35. Это означает, что страны с более низкой среднегодовой температурой имеют более перспективные условия для экономи-

ческого развития. Следует отметить, что между этими переменными до настоящего времени не установлено наличие ни функциональной, ни статистической связи. Да их, по видимому, и не существует. Обращает на себя внимание тот факт, что коэффициент корреляции между ВВП государства и его среднегодовой температурой отрицателен.

Далее проводилось выделение скрытых факторов, устанавливающих связь между, кажущимися хаотичными, наблюдаемыми показателями. Для этого определялись матрицы A и F – матрицы факторных нагрузок и значений соответствующих факторов. Для большей наглядности рассматривались лишь два фактора Φ_1 и Φ_2 . Эти факторы являются общими (фактор считается общим, если хотя бы две его нагрузки отличны от нуля).

В таблице 3 представлена воспроизведенная корреляционная матрица, полученная путём факторного отображения.

Табл.3.

Воспроизведённая матрица

Переменная	Vozr	Plos	Nas	VVP	Temp
Vozr	0,27	0,28	0,41	-0,30	0,30
Plos	0,28	0,72	0,64	0,03	-0,08
Nas	0,41	0,64	0,74	-0,29	0,27
VVP	-0,30	0,03	-0,29	0,64	-0,67
Temp	0,30	-0,08	0,27	-0,67	0,71

Сумма ее диагональных элементов, или общность, для двух выбранных факторов составляет величину $\approx 3,04$ при полной дисперсии равной 5 (пять показателей). Это означает, что с помощью двух рассматриваемых факторов можно воспроизвести более 60% информации об исследуемых связях и явлениях. Причём,

на первый фактор приходится более 37,6% общей дисперсии, а на второй – более 23,6%.

Важная информация о связях факторов и наблюдаемых переменных может быть получена с помощью геометрической интерпретации факторного анализа. В частности, при двух факторах можно перейти в двумерное пространство общих факторов. Для данной задачи эти результаты представлены на рисунке 1.

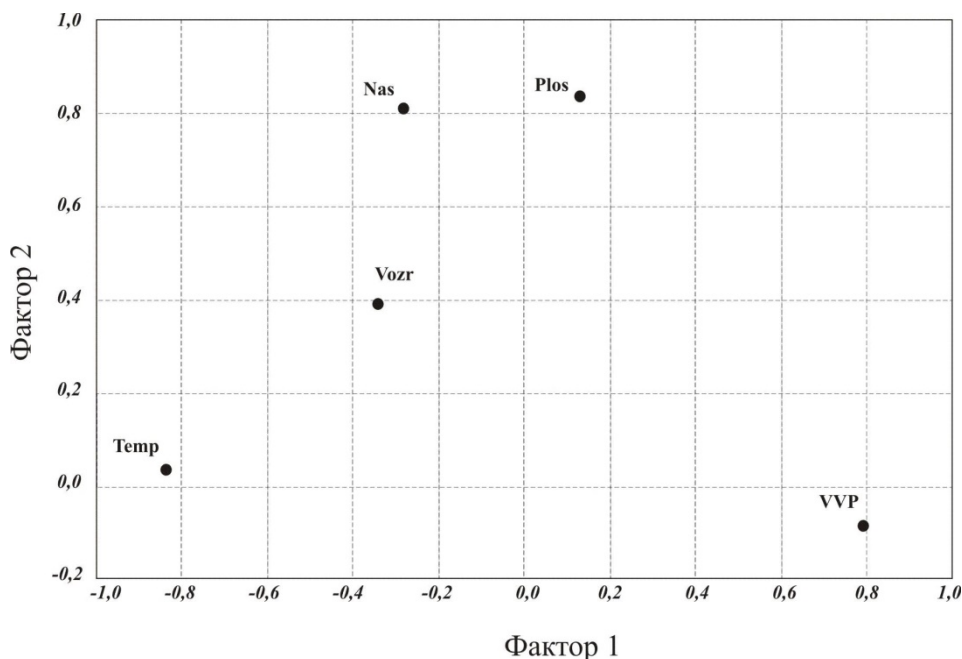


Рис. 1. Двумерное пространство общих факторов.

По осям абсцисс и ординат отложены нормированные факторы Φ_1 и Φ_2 . Значения каждой переменной в факторном пространстве соответствуют точке, координатами которой являются соответствующие факторные нагрузки. Следует отметить, что расположение переменных в факторном пространстве может быть так же представлено векторами с соответствующими проекциями. Из анализа представленной двумерной факторной структуры следуют важ-

ные выводы. Такой значимый показатель государства, как ВВП, имеет максимальную проекцию на первый фактор со значением 0,73 и ее знак положителен. А проекция на этот же фактор среднегодовой температуры – отрицательна и равна -0,78. Проекция этих показателей на второй фактор – малы, ими можно пренебречь. Таким образом, существует некий скрытый фактор, для данного случая это Φ_1 , неподдающийся описанию, который определяет корреляционную

связь между этими величинами со значимым коэффициентом корреляции равным $-0,67$ (таблица 3).

Этот результат можно трактовать следующим образом. Страны с более низкой среднегодовой температурой имеют более высокий уровень экономического развития, и этот результат подтверждается реальными фактами. В частности, к странам с наиболее высоким ВВП и высоким жизненным уровнем относятся Канада, Норвегия, Исландия, скандинавские страны. Ещё пример – при освоении Северной Америки в XVI-XVII веках, образовавшиеся на ее территории страны Мексика, США и Канада имеют различный уровень экономического развития, который в какой-то мере коррелирует с отрицательным коэффициентом корреляции со среднегодовой температурой. Развитые в свое время Древний Египет, Древний Рим и Древняя Греция развивались в благоприятных климатических условиях, но спустя тысячелетия уступили свои лидирующие позиции, что нельзя сказать о странах с жесткими климатическими условиями.

ВВП и такие геополитические показатели, как возраст государства и количество населения так же имеют противоположные знаки проекций на ось первого фактора. Это подтверждается и соответствующими отрицательными коэффициентами корреляции воспроизведенной матрицы (таблица 3). Однако ВВП и площадь страны связаны небольшими, но положительными коэффициентами корреляции.

Интересным является вопрос раскрытия сущности выделенных факторов, хотя сделать это довольно затруднительно, а порой и невозможно. Ситуация упрощается, если пара переменных имеет одну из факторных проекций близкую к нулю, как, например, среднегодовая температура и ВВП. Проекция этих показателей

на второй фактор – малы, как отмечалось выше. При этих условиях можно интерпретировать сущность первого фактора следующим образом. Более жесткие климатические условия формируют определенные качества членов общества, особый менталитет, более рациональное отношение к воспроизводимому материальному продукту и способам его получения, что, в конечном счете, приводит к более высокому качеству жизни.

Дальнейший анализ полученных корреляционных зависимостей между наблюдаемыми геополитическими характеристиками и влияния скрытых факторов, обуславливающих связи между ними, может иметь значительное продолжение. В частности, нельзя отрицать тот факт, что существуют некоторые оптимальные значения геополитических характеристик, при которых возможно достижение максимального ВВП. Ими могут быть оптимальная площадь государства, количество населения, возможно протяженность морских границ, уровень образованности нации и др. Этот вопрос, по-видимому, станет предметом исследований социологов, политологов и экономистов.

ВЫВОДЫ

В заключение следует отметить, что в данной работе предложен и продемонстрирован комбинированный подход применения методов системного анализа для оценки эффективности функционирования системы на примере государства, базирующийся на многомерном статистическом анализе, и факторном анализе, в частности, который может быть использован при проведении исследований в области государственного развития и управления.

REFERENCES

1. Francis Fukuyama. The End of History and the Last man. – N-Y: Free Press. – 2005. – 612p.
2. Socialno-ekonomicheskay geografia zarubejnogo mira. // Red.V.Volskogo. – M.:CRON-PRESS. – 1998. – 592 s.
3. LipetsY., Pyulyarkn V., Shlihter S. Geographiya mirovogo hozyaistva. – M.: Gumanit. – 2009. – 400 s.
4. Kant I. On history. – Indianapolis: Bobbs-Merrill. – 1963. 154 p.
5. Shamrov's'kyy O. D. Sistemniy analiz: matematychni metody ta zastosyvannya. – Lviv: Magnolia. – 2015. – 273 s.
6. Anderson T. W. An introduction to multivariate statistical analysis. – N.Y. – 2008. – 131 s.
7. Überla K. Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler. Springer-Verlag Berlin Heidelberg New-York. – 2001. – 198 s.
8. Spitsnadel' V.N. Osnovy sistemnogo analiza. M.: Biznes-Pressa. – 2000. – 326 s.
9. Anfilatov V.S. Sistemnyy analiz v upravlenii. – M.: Finansy i statistika. – 2009. – 368 s.
10. Popov V.B. Sistemnyy analiz v upravlenii: Uchebnoye posobiye / V.B. Popov. – M.: Finansy i statistika. – 2009. – 288 s.

Рецензент: д.т.н., проф. Фісун М. Т.
Чорноморський державний університет ім. Петра Могили