

УДК 656.071

БОСОВ А.А., д.т.н., профессор (ДНУЗТ)
ЛОЗА П.А., к.т.н., доцент (ДНУЗТ)

Построение индекса произвольного процесса

Bosov A., Dr. Eng., Professor (DNURT)
Loza P., PhD in Eng., Associated Professor (DNURT)

Creation of an index of arbitrary process

Введение

Многие процессы или явления, протекающие во времени, характеризуются несколькими показателями, и тогда возникает задача, с помощью линейных преобразований, определить такие показатели, которые между собой не коррелировали, но незначительное их число с достаточной степенью точности описывало бы исходный процесс.

Данная задача решается с помощью метода главных компонент. В предлагаемой работе осуществляется объединение

двух методов: метода главных компонент и метода анализа иерархий, что позволяет избавиться от субъективизма экспертов в методе анализа иерархий.

Основной материал

Пусть в любой момент времени некоторый процесс характеризуется набором показателей $x_i, i = \overline{1, n}$.

Тогда возникает задача введения некоторого индекса, который бы с опреде-

ленной степенью точности отражал бы поведение процесса во времени.

Если в моменты времени $t_1, t_2, \dots, t_m$ фиксировались значения показателей, то располагаем матрицей значений показателей. Данная матрица имеет n строк и m столбцов, а x_{ij} будет представлять собой значение показателя x_i в момент времени t_j .

Процедура построения индекса будет состоять из двух этапов.

На первом этапе с помощью линейных преобразований переменных $x_i, i = \overline{1, n}$ определяются новые показатели $z_i, i = \overline{1, n}$ такие, что они между собой не коррелируют [1].

Другими словами, основными являются соотношения

$$x_i = \sum_{r=1}^n w_{ir} z_r, \quad i = \overline{1, n} \quad (1)$$

или в матричной форме

$$x = wz \quad (2)$$

Как показано в работе [2] первоначально приходим к переменным $y_1, y_2, \dots, y_n$, которые удовлетворяют равенствам

$$y = U'x, \quad x = Uy,$$

где U – ортогональная матрица.

Если u_1 первый столбец матрицы U , то его выбираем так, чтобы дисперсия y_1 была бы максимальной. После чего выбираем второй столбец u_2 , так чтобы дисперсия y_2 была бы максимальной, и выполнялось условие некоррелированности y_2 с y_1 .

Аналогично процедура продляется далее, и если мы работаем с y_r , то накладывается условие некоррелированности y_r с $y_1, y_2, \dots, y_{r-1}$.

Если λ_r дисперсия y_r , а A ковариационная матрица x , то учитывая

$$y_r = u_r'x$$

получим

$$\lambda_r = u_r' A u_r$$

Поскольку y_i некоррелированная, то имеет место

$$u_r' A u_s = 0, \quad r \neq s,$$

а это означает, что матрица

$$\Lambda = U' A U$$

является диагональной с элементами $\lambda_1, \lambda_2, \dots, \lambda_n$.

Более того имеет место

$$A u_r = \lambda_r u_r, \quad r = \overline{1, n},$$

тогда нормируя y_r получаем r -тую главную компоненту z_r .

Соотношение (2) теперь можно записать в виде

$$x = U \Lambda^{1/2} z,$$

т.е. матрица W будет следующей

$$W = U \Lambda^{1/2}.$$

Если вместо ковариационной матрицы A использовать корреляционную, то тогда имеет место

$$\sum_{r=1}^n \lambda_r = n.$$

Выбирая первые k главных компонент, получим, что они описывают следующую долю разброса исходных показателей

$$\frac{1}{n} \sum_{r=1}^k \lambda_r,$$

$$w_i = \frac{a_{1n}}{a_{1i}}, i = \overline{1, n}.$$

тогда вместо x вводя $z_1, z_2, \dots, z_k$ главных компонент с заданной точностью описываем изучаемый процесс.

На втором этапе производим линейную свертку k главных компонент по методу анализа иерархий [3].

Существенным элементом данного метода является построение матрицы попарных сравнений A , элементы которой должны обладать свойствами:

c1. все $a_{ij} > 0$;

c2. $a_{ij} = 1 / a_{ji}$;

c3. $a_{ik} a_{kj} = a_{ij}$ (свойство совместности);

c4. максимальное собственное значение матрицы A должно быть равно n .

В работе [4] показано, что если задана первая строка матрицы A т.е. $1, a_{12}, a_{13}, \dots, a_{1n}$, то положив

$$a_{ij} = \frac{a_{1j}}{a_{1i}}, i = \overline{2, n}, j = \overline{1, n},$$

все четыре свойства матрицы A будут выполнены, а компоненты собственного вектора с максимальным собственным значением вычисляются по формуле

Если $z_1, z_2, \dots, z_k$ главные компоненты, определяемые по методу главных компонент, а $s_1, s_2, \dots, s_k$ дисперсии разброса описываемые главными компонентами, то положив первую строку матрицы A по формуле

$$a_{11} = 1; a_{12} = \frac{s_2}{s_1}, \dots, a_{1i} = \frac{s_i}{s_1}, i = \overline{2, k}$$

получаем возможность вычисления $w_i, i = \overline{1, n}$, и после нормировки индекс процесса будет равным

$$z = \sum_{i=1}^k w_i z_i. \quad (3)$$

Таким образом, комбинируя метод главных компонент и метод анализа иерархий, получаем индекс процесса (3) как некоторую функцию исходных показателей.

В качестве примера рассмотрим индекс социальных последствий работы железной дороги. Показатели их значения за период с 2001 по 2012 годы представлены в Таблице 1.

Таблица 1
Справка о выделении средств железной дорогой на социальные нужды (тыс.грн.)

№ п/п	Наименование	год											
		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
1	Численность работников, чел.	67618	66066	66045	65047	64922	64837	63375	63757	62298	61122	60445	58654
2	Фонд оплаты труда, тыс. грн.	435020	549822	656933	766582	903139	1140860	1403447	1835556	1883050	2098018	2371558	2646323
3	Средняя заработная плата, грн.	536	694	829	982	1159	1466	1845	2399	2519	2860	3270	3760
4	Производительность труда, тыс. т-км на 1 работн.	1163	1154	1279	1320	1271	1329	1546	1516	1252	1388	1541	1556
5	Материальная помощь не работающим, грн	9946	22602	21751	29403	30392	18255	12308	17274	14504	14248	16515	27059
6	Материальная помощь на оздоровление, грн	7020	7823	11027	12628	17930	22030	28608	37712	44165	54206	58634	64990
7	Оказание мат. помощи при выходе на пенсию	6660	5596	2930	4414	5250	8769	14245	18829	21472	21490	76277	23869
8	Финансовая поддержка оздоровит. спортивно-	2278	6662	5970	6332	11464	12739	20600	18947	20132	22886	24860	28854

ОРГАНІЗАЦІЯ ТА УПРАВЛІННЯ ПРОЦЕСОМ ПЕРЕВЕЗЕНЬ

	культурным заведениям, детским жел. дорогам												
9	Мат. помощь медицинским заведениям	3960	3987	3490	11418	9834	8634	9945	10001	13246	13640	13965	12744
10	Мат. помощь учебным заведениям	1832	2737	7798	19437	2325	1668	425	94	37	148	0	0
11	Бесплатные билеты для личных потребностей работающих	0	0	0	0	10237	10802	11170	13728	18124	18519	24729	34060
12	Бесплатные билеты для личных потребностей неработающих	0	0	0	0	3634	3415	3600	3725	4852	5257	6764	4062
13	Отчисления профсоюзу, в том числе	6316	5131	4996	8997	11343	20750	21362	25917	24281	27255	29345	31490
14	-отчисления профсоюзным орг. на провед. культ.-масс. и спортивной работы	546	1641	2882	4337	5382	6132	6192	7800	7195	7695	8282	8295
15	-удешевление стоимости путевок в оздоровит. учреждения железной дороги	1284	2190	1196	3847	4463	8003	12412	13723	13288	15360	16671	17991
16	- санаторные путевки	1150	1300	918	813	1498	1738	2758	3000	3500	4200	4392	5204

Как следует из таблицы 1, число показателей $n=16$, а число наблюдений (годов) $m=12$.

При числе наблюдений равных 12, критический коэффициент корреляции $K_r = 0.58$ (см. [5] таблица XII)

Задавшись долей описываемого разброса данных $\alpha = 0.9$, по методу главных компонент определяем число главных ком-

понент, которые описывают 16 показателей.

На рис. 1 представлены три главные компоненты.

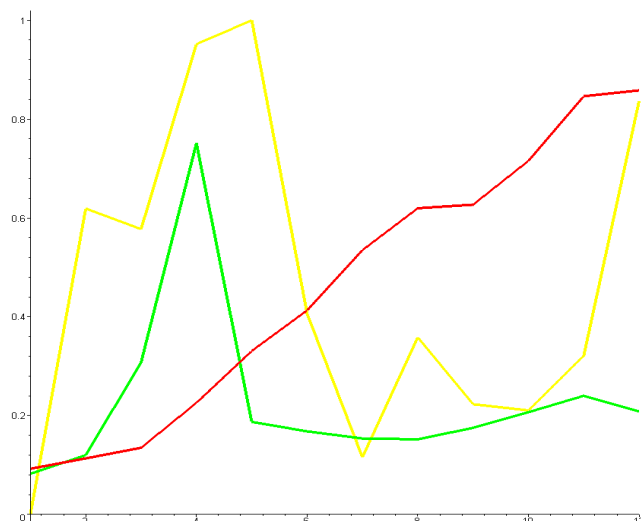


Рис. 1. Главные компоненты (z_1, z_2, z_3)

Компонента z_1 описывает $s_1 = 0.765$ разброса данных, а z_2 имеет $s_2 = 0.091$, компонента z_3 имеет $s_3 = 0.063$.

Далее используем метод анализа иерархий, полагая

$$a_{11} = 1; a_{12} = \frac{s_2}{s_1}; a_{13} = \frac{s_3}{s_1},$$

которым соответствуют следующие веса главных компонент

$$w = [0.046, 0.387, 0.566]$$

С помощью данных весов строим индекс социальных последствий по формуле

$$z = w_1 z_1 + w_2 z_2 + w_3 z_3$$

График данного индекса представлен на рис.2.

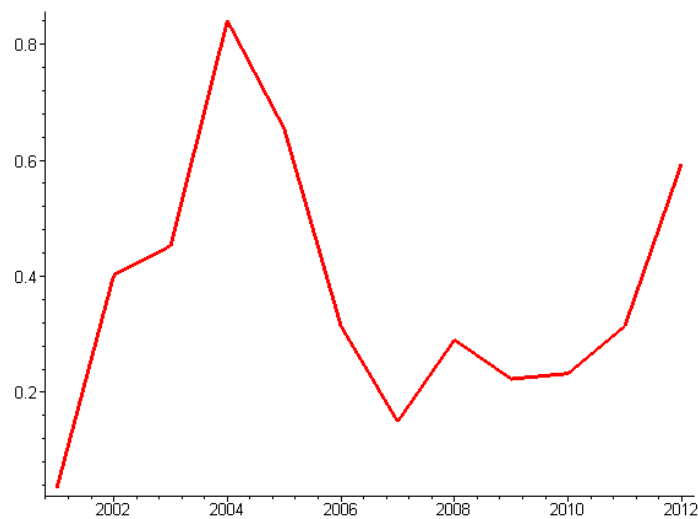


Рис. 2. Индекс социальных последствий работы железной дороги в период с 2001 по 2012 годы

Как следует из рис.2, начиная с 2004 года, имеет место падение индекса по 2007 год, а с 2007 года индекс постепенно возрастает. Хотя с 2008 по 2009 год он незначительно упал, а далее наблюдается его рост.

Отмеченные колебания вполне можно объяснить ситуациями в стране, которые происходили в рассматриваемый период времени.

Выводы

1. Для дискретных процессов предложен метод построения его индекса. Данный метод представляет собой комбинацию метода главных компонент и метода анализа иерархий. В предлагаемом методе исключается субъективизм экспертов, с помощью

использования дисперсий для главных компонент.

2. Предложенный метод используется для расчета индекса социальных последствий работы железной дороги.

3. Данный метод может быть использован для оценки мероприятий, посвященных социальной сфере и определения такого набора мероприятий, которые укладываются в заданные объемы финансирования и максимально улучшающих социальную сферу.

Список литературы:

1. Pearson K. On lines and planes of closest fit to system of points in space. Phil. Mag.2, 6th series, 1901, p.557-572.

2. Лоули Д. Факторный анализ как статистический метод / Д. Лоули, А. Максвелл. – М.: Мир, 1967. – 144с.

3. Саати Т. Принятие решений: метод анализа иерархий. – М.: Радио и связь, 1993. – 320с.

4. Ногин В.Д. Упрощенный вариант метода анализа иерархий на основе нелинейной свертки критериев. – С.-Петербург, 2004. – 16 с.

5. Плохинский Н.А. Биометрия. – Из-во Московского университета, 1970, – 367с.

Spisok literatury:

1. Pearson K. On lines and planes of closest fit to system of points in space. Phil. Mag.2, 6th series, 1901, p.557-572.

2. Louli D. Faktornyiy analiz kak statisticheskiy metod / D. Louli, A. Maksvell. – М.: Mir, 1967. – 144с.

3. Saati T. Prinyatie resheniy: metod analiza ierarhiy. – М.: Radio i svyaz, 1993. – 320с.

4. Nogin V.D. Uproschennyiy variant metoda analiza ierarhiy na osnove nelineynoy svertki kriteriev. – S.-Peterburg, 2004. – 16 s.

5. Plohinskiy N.A. Biometriya. – Iz-vo Moskovskogo universiteta, 1970, – 367s.

Аннотации:

В предлагаемой работе осуществляется объединение двух методов: метода главных компонент и метода анализа иерархий.

Предложен алгоритм построения индекса по наблюдаемым показателям.

При построении индекса произвольного процесса удалось избавиться от субъективизма экспертов используя метод анализа иерархий.

Если процесс является управляемым, то получаем возможность оценки, как принимавшихся решений, так и определения рациональных решений в определенном смысле на исследуемый процесс.

Ключевые слова: метод главных компонент, анализ иерархий, индекс процесса.

У пропонованій роботі здійснюється об'єднання двох методів: методу головних компонент і методу аналізу ієрархій.

Запропоновано алгоритм побудови індексу по спостережуваних показниками.

При побудові індексу довільного процесу вдалося позбутися суб'єктивізму експертів використовуючи метод аналізу ієрархій.

Якщо процес є керованим, то отримуємо можливість оцінки як рішень, що приймалися, так і визначення раціональних рішень у певному сенсі на досліджуваний процес.

Ключові слова: метод головних компонент, аналіз ієрархій, індекс процесу.

In offered work is carried out join of two methods: method of the principal components and method of the analysis of hierarchies.

The algorithm of creation of an index on observed indexes is offered.

At creation of an index of the arbitrary process, succeeded to get rid of subjectivity of experts, using a method of the analysis of hierarchies.

If process is operated, we have an assessment opportunity, both made decisions, and definition of rational decisions in a sense on studied process.

Keywords: method of the principal components, analysis of hierarchies, process index.