

Липецкий государственный технический университет

Факультет автоматизации и информатики

Кафедра автоматизированных систем управления

ЛАБОРАТОРНАЯ РАБОТА №3

СТАТИСТИЧЕСКИЕ МЕТОДЫ В ПРИКЛАДНЫХ ЗАДАЧАХ

ФАКТОРНЫЙ АНАЛИЗ

Студент

подпись

Станиславчук С.М.
Коновалов К.А.

Группа АС-21-1

Руководитель

Доцент

подпись

Рыжкова Д.В.

Липецк 2023 г.

Задание кафедры

1. По исходным данным согласно номерам вариантов, за исключением переменных из таблицы 1, провести факторный анализ, не совершая поворот факторов, методом главных компонент. Получить матрицу парных коэффициентов корреляции. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.
2. На основе полученного решения произвести поворот факторов методом Варимакс нормализованный. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.
3. На основе решения, полученного в Задании 1, произвести поворот факторов методом Варимакс исходный. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.
4. Провести факторный анализ центроидным методом. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.

Ход выполнения лабораторной работы

	1 Y1	2 X4	3 X6	4 X8	5 X9	6 X10	7 X11	8 X12	9 X13	10 X14	11 X15	12 X16	13 X17
19	8.15	1.12	0.77	0.43	0.29	0.18	64044	103.73	107636	6.23	67.16	14.76	13.44
20	8.72	0.99	1.08	0.14	1.87	2.24	34328	38.95	67592	4.25	50.4	7.56	13.69
21	6.64	0.58	0.93	0.29	0.47	0.3	58424	81.32	99812	5.38	70.89	8.64	16.66
22	8.1	1.03	0.1	0.43	0.34	0.15	83240	67.75	75680	5.88	72	8.64	15.06
23	5.52	1.24	0.11	0.17	0.27	0.17	6462	59.66	44196	9.27	97.2	9	20.09
24	9.37	0.89	1.44	0.21	0.8	2.3	114896	107.81	20898.5	4.36	80.28	14.76	15.91
25	13.17	0.68	0.48	0.42	0.97	0.31	21791.47	512.62	28946	10.31	51.48	10.08	18.27
26	6.67	1.03	1.24	1.19	1.39	0.44	83568	53.53	74687.88	4.72	105.12	14.76	14.44
27	5.68	0.73	0.77	1.87	0.16	0.18	68976	80.83	8631.25	4.18	128.52	10.38	22.88
28	5.19	0.73	0.93	0.15	0.15	0.39	67663.88	59.42	31314	3.13	94.68	14.76	15.5
29	10.02	0.85	0.13	0.03	1.15	2.6	34428	36.96	64752	4.02	85.32	20.52	19.35
30	8.16	1.03	1.73	0.24	0.21	0.45	127256	91.88	8206.25	5.2	76.32	14.46	16.95
31	3.78	0.47	0.77	0.93	0.89	0.45	6265	17.16	44676	2.72	153	24.88	30.53
32	6.45	0.56	0.16	0.13	1.15	2.25	33192	27.29	65188	3.12	107.34	11.16	17.78
33	10.38	0.89	0.74	0.27	0.13	0.49	127983.88	184.33	22697	10.38	90.72	6.45	22.09
34	7.65	0.99	1.95	0.17	0.33	0.14	41368	58.42	68104	5.65	82.44	9.72	18.29
35	8.77	1.95	0.58	0.24	0.64	0.18	33556	59.31	65616	6.67	79.12	3.24	26.05
36	7	1.03	1.77	0.19	0.93	0.29	124560	49.87	127344	5.93	120.96	6.45	26.2
37	11.06	0.01	0.7	0.29	0.14	0.5	110548	391.27	7919	11.89	84.6	5.4	17.26
38	9.02	0.02	0.74	0.25	0.13	0.26	95968	258.61	14314.75	8.3	85.32	6.12	18.95
39	13.28	0.6	1.15	0.36	0.16	0.21	21182.5	75.14	9277.13	0.18	101.52	8.64	19.66
40	9.27	0.97	1.19	0.17	0.49	0.49	53284	123.16	122072	8.88	107.34	11.94	16.97
41	6.73	1.12	1.03	0.23	1.8	0.28	6338	37.21	85792	5.82	85.32	7.92	14.63
42	6.72	1.77	1.08	0.17	0.15	0.15	44460	53.37	79631.88	4.8	131.76	10.08	22.17
43	9.44	0.93	0.13	0.24	1.99	2.03	6555	32.87	40516	5.01	116.64	18.72	22.62
44	7.21	1.12	0.74	0.26	0.1	0.4	44340	45.87	72580	4.12	138.24	13.68	26.44
45	5.39	0.74	0.99	0.13	0.17	0.27	35888	48.41	72296	5.1	156.96	16.56	22.26
46	5.61	0.47	0.64	0.28	0.21	0.38	3947.5	13.58	7612	3.47	137.52	14.76	19.13
47	5.59	1.12	1.87	0.34	1.65	0.45	59356	63.69	49072	4.22	134.06	7.92	18.28
48	6.57	0.58	1.12	0.26	0.64	0.14	75304	104.55	22896	5.01	155.52	18.36	28.23
49	6.54	0.64	0.08	4.47	0.2	0.43	68712	223.7	10949.73	11.38	48.6	8.28	12.44
50	4.2	0.6	0.17	0.25	1.92	0.14	3351	25.88	6235	7.67	42.84	14.04	11.64
51	5.19	1.53	0.03	2.13	1.65	0.17	6369	29.52	84000	4.68	143.64	16.79	8.62
52	18	1.73	0.02	0.27	0.09	0.18	90336	41.69	65050	4.3	145.8	11.16	20.1

Рисунок 1 – Данные для анализа

Краткая теоретическая справка

Основное понятие факторного анализа (фактор) определяется следующим образом: *факторы* - гипотетические непосредственно не измеряемые, скрытые (латентные) переменные в той или иной мере связанные с измеряемыми характеристиками – проявлениями этих факторов.

Идея факторного анализа основана на предположении, что имеется ряд величин, не известных исследователю, которые заставляют проявляться различные соотношения между переменными.

Под факторным анализом понимают методы выявления гипотетических (ненаблюдаемых) факторов, призванных объяснить корреляционную матрицу количественных наблюдаемых переменных.

Принято считать статистический анализ такого рода успешным, если большое число переменных удалось объяснить малым числом факторов.

Конечная цель статистического исследования, проводимого с привлечением факторного анализа, как правило, состоит в выявлении и

интерпретации скрытых общих факторов с одновременным стремлением минимизировать их число и степень зависимости от характерных факторов.

Факторный анализ позволяет выявить зависимость между явлениями, обнаружить скрытую основу нескольких явлений, ответить на вопрос, почему связаны явления.

Как метод статистического исследования факторный анализ включает следующие основные этапы:

- 1) Формирование цели: исследовательские (выявление факторов и их анализ); прикладные (построение агрегированных характеристик для прогнозирования и управления).
- 2) Выбор совокупности признаков и объектов.
- 3) Получение исходной факторной структуры.
- 4) Корректировка факторной структуры, исходя из целей исследования.
- 5) Выявление факторов второго порядка.
- 6) Интерпретация и использование результатов.

Задание 1. По исходным данным согласно номерам вариантов, за исключением переменных из таблицы 1, провести факторный анализ, не совершая поворот факторов, методом главных компонент. Получить матрицу парных коэффициентов корреляции. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.

Переменная	Средние	Ст. откл.	Y1	X4	X6	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17
Y1	7.91	2.60	1.000000	0.091458	0.032344	-0.115588	-0.065245	0.057981	0.274654	0.447442	-0.119111	0.201132	-0.102520	-0.259043	0.019287
X4	0.96	0.41	0.091458	1.000000	-0.051222	-0.090179	0.115961	-0.186331	-0.068345	-0.360488	0.349009	-0.012560	0.213446	-0.147227	-0.006900
X6	0.89	0.58	0.032344	-0.051222	1.000000	-0.250922	0.024004	-0.046128	0.229372	-0.004290	0.084638	-0.170735	0.283167	-0.100940	0.218011
X8	0.47	0.71	-0.115588	-0.090179	-0.250922	1.000000	-0.063897	-0.139096	0.004619	0.156207	-0.144978	0.199762	-0.156080	0.068959	-0.258754
X9	0.75	0.61	-0.065245	0.115961	0.024004	-0.063897	1.000000	0.298519	-0.253571	-0.100674	-0.052211	0.004488	-0.010200	0.079547	-0.048574
X10	0.53	0.64	0.057981	-0.186331	-0.046128	-0.139096	0.298519	1.000000	-0.049110	-0.144937	0.060830	-0.267441	-0.146348	0.299247	-0.078989
X11	52314.36	38332.40	0.274654	-0.068345	0.229372	0.004619	-0.253571	-0.049110	1.000000	0.310252	-0.046024	0.144753	-0.079948	-0.274606	-0.038569
X12	92.16	93.26	0.447442	-0.360488	-0.004290	0.156207	-0.100674	-0.144937	0.310252	1.000000	-0.370344	0.623973	-0.287422	-0.305538	-0.052185
X13	48853.30	34065.36	-0.119111	0.349009	0.084638	-0.144978	-0.052211	0.060830	-0.046024	-0.370344	1.000000	-0.261376	0.067999	0.002895	-0.087319
X14	5.93	2.45	0.201132	-0.012560	-0.170735	0.199762	0.004488	-0.267441	0.144753	0.623973	-0.261376	1.000000	-0.257093	-0.317239	-0.142839
X15	108.52	35.97	-0.102520	0.213446	0.283167	-0.156080	-0.010200	-0.146348	-0.079948	-0.287422	0.067999	-0.257093	1.000000	0.208907	0.547159
X16	12.85	5.53	-0.259043	-0.147227	-0.100940	0.068959	0.079547	0.299247	-0.274606	-0.305538	0.002895	-0.317239	0.208907	1.000000	0.074557
X17	19.57	4.74	0.019287	-0.006900	0.218011	-0.258754	-0.048574	-0.078989	-0.038569	-0.052185	-0.087319	-0.142839	0.547159	0.074557	1.000000

Рисунок 2 – Корреляционная матрица

Проанализируем корреляционную матрицу. Заметим сильную корреляционную связь между переменными “фондовооруженность труда “

X14 и “среднегодовая стоимость ОПФ” X12. Поэтому не будем удивляться, увидев их сильную факторную связь для Фактора 1.

Фактор.нагрузки (Без вращ.) (Spreadsheet1) Выделение: Главные компоненты (Отмечены нагрузки > ,7000000)				
Перемен.	Фактор 1	Фактор 2	Фактор 3	
Y1	0,461413	-0,332865	-0,021600	
X4	-0,276690	-0,213529	0,734499	
X6	-0,154843	-0,599185	-0,208702	
X8	0,300740	0,431605	0,095768	
X9	-0,198396	0,282268	-0,105072	
X10	-0,246253	0,381007	-0,391446	
X11	0,430786	-0,415936	-0,007590	
X12	0,846310	-0,126981	-0,259119	
X13	-0,428231	-0,065950	0,568861	
X14	0,735706	0,006453	0,121827	
X15	-0,528038	-0,541192	-0,154523	
X16	-0,500589	0,375067	-0,418362	
X17	-0,300761	-0,589808	-0,400166	
Общ.дис.	2,750214	1,901496	1,521636	
Доля общ	0,211555	0,146269	0,117049	

Рисунок 3 – Матрица факторных нагрузок

Проанализируем таблицу с текущими факторными нагрузками. Факторные нагрузки могут интерпретироваться как корреляции между соответствующими переменными и факторами – чем выше нагрузка по модулю, тем больше близость фактора к исходной переменной.

Фактор 1. Данное значение факторной нагрузки может говорить о том, что существует сильная положительная связь между среднегодовая стоимость ОПФ и фондовооруженность труда с фактором 1. Возможно, что данный фактор может связываться с эффективностью использования ресурсов производства, поскольку обе оценки имеют отношение к экономическим аспектам производства.

Фактор 2. Имеет не очень сильные связи с переменными, однако стоит отметить отрицательную связь “удельный вес покупных изделий” X6, “оборачиваемость нормированных оборотных средств” X15 и “непроизводственные расходы” X17.

Фактор 3. Данный фактор имеет сильную связь лишь с одной переменной – “трудоемкость единицы продукции” X4, на этой основе можно предположить, что это довольно важный фактор, который связан, такими как эффективность производственных процессов, сложность выполняемых задач, квалификация работников и т.д.

Остальные факторные нагрузки не сильно связаны или вообще нулевые

Задание 2. На основе полученного решения произвести поворот факторов методом Варимакс нормализованный. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.

Фактор.нагрузки (Варимакс нормализ.) (Spreadsheet1)					
Выделение: Главные компоненты					
(Отмечены нагрузки >,700000)					
Перемен.	Фактор 1	Фактор 2	Фактор 3		
Y1	-0,532418	0,043212	0,197056		
X4	-0,157002	0,026383	-0,797681		
X6	-0,185105	0,625342	0,035183		
X8	0,012118	-0,522694	0,111996		
X9	0,349106	-0,084731	0,031977		
X10	0,533533	-0,030770	0,270985		
X11	-0,564815	0,118203	0,160152		
X12	-0,614537	-0,217744	0,611915		
X13	0,095080	0,047305	-0,707146		
X14	-0,571865	-0,411997	0,243672		
X15	0,105885	0,742458	-0,182030		
X16	0,724225	0,108838	0,173042		
X17	-0,012904	0,761802	0,133998		
Общ. дис.	2,317683	2,051747	1,803917		
Доля общ.	0,178283	0,157827	0,138763		

Рисунок 4 – Матрица факторных нагрузок

Варимакс нормализованный - это метод варимакс, который используется для поворота факторов с нормализацией факторных нагрузок. Это означает, что масштабы факторных нагрузок будут одинаковыми, что облегчит интерпретацию полученных факторов.

Как можем заметить, здесь уже участвуют в сильной связи совсем другие переменные.

Фактор 1. Сильная факторная связь между оборачиваемостью ненормированных оборотных средств и фактором 1 может указывать на то, что этот фактор связан с эффективностью управления запасами и/или операционной эффективностью предприятия в целом. Высокая оборачиваемость ненормированных оборотных средств может говорить о том, что предприятие оперирует своими ресурсами более эффективно, что может привести к увеличению прибыли.

Фактор 2. Оборачиваемость нормированных оборотных средств и непроизводственные расходы имеет сильную связь с фактором 2. Это может говорить о том, что фактор 2 может быть связан с управленческой эффективностью в использовании оборотных средств и контролем непроизводственных расходов. Возможно, это связано с тем, как организация управляет своими финансами и как ей удастся управлять своими операционными расходами.

Фактор 3. Трудоемкость единицы продукции и среднегодовой фонд заработной платы имеют сильную обратную связь с фактором 3. Если трудоемкость единицы продукции и среднегодовой фонд заработной платы имеют сильную обратную связь с фактором 3, то можно предположить, что данный фактор связан с эффективностью использования трудовых ресурсов. Возможно, этот фактор отражает производственные процессы и технологии, которые позволяют снизить трудоемкость и повысить производительность труда, что в свою очередь может снизить затраты на заработную плату.

Задание 3. На основе решения Задания 1 произвести поворот факторов методом Варимакс исходный. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.

Перемен.	Фактор.нагрузки (Вар.исходн.) (Spreadsheet1) Выделение: Главные компоненты (Отмечены нагрузки >,700000)					
	Фактор 1	Фактор 2	Фактор 3			
Y1	0,548951	0,098346	0,114665			
X4	0,035431	0,012847	-0,812539			
X6	0,127866	0,640446	-0,005830			
X8	0,054216	-0,517417	0,123445			
X9	-0,331005	-0,115014	0,085347			
X10	-0,482789	-0,069805	0,347964			
X11	0,568314	0,174648	0,071792			
X12	0,715325	-0,139476	0,518038			
X13	-0,201537	0,013855	-0,685948			
X14	0,637954	-0,349483	0,164388			
X15	-0,201576	0,722954	-0,179708			
X16	-0,698171	0,048367	0,276603			
X17	-0,040236	0,764004	0,114644			
Общ. дис.	2,476508	1,986820	1,710018			
Доля общ.	0,190501	0,152832	0,131540			

Рисунок 5 – Матрица факторных нагрузок

Варимакс исходный - это метод варимакс, который используется для поворота факторов без нормализации факторных нагрузок. Это означает, что масштабы факторных нагрузок могут быть разными, что может затруднить интерпретацию полученных факторов.

Фактор 1. Наблюдается сильная связь с среднегодовой стоимостью ОПФ X12. Это может указывать на то, что данный фактор связан с экономическими аспектами, связанными с капитальными вложениями или инвестициями. В данном случае, так как среднегодовая стоимость ОПФ имеет высокую факторную нагрузку на фактор 1, можно предположить, что этот фактор связан с финансовыми аспектами, связанными с капиталовложениями, например, с общим уровнем инвестиций, эффективностью использования капитальных вложений и т.д.

Фактор 2. Аналогично с предыдущим методом Варимакс нормализованный.

Фактор 3. Это означает, что чем выше значение фактора 3, тем меньше значение переменной X4, то есть трудоемкость единицы продукции снижается. Это может быть связано с тем, что фактор 3 может отражать процессы автоматизации производства, внедрение новых технологий, которые позволяют сократить трудоемкость.

Задание 4. Провести факторный анализ центроидным методом. Получить первые три общие фактора и дать им экономическую интерпретацию по матрице факторных нагрузок.

Перемен.	Фактор.нагрузки (Без вращ.) (Spreadsheet1) Выделение: Главные факторы (Центроид) (Отмечены нагрузки >,700000)				
	Фактор 1	Фактор 2	Фактор 3		
Y1	0,352716	0,106274	-0,097710		
X4	-0,095950	-0,157482	-0,316098		
X6	0,285242	-0,370648	-0,164186		
X8	-0,039668	0,363670	0,151927		
X9	-0,227555	-0,103194	0,180743		
X10	-0,287335	-0,047645	0,055429		
X11	0,437374	0,119072	-0,318302		
X12	0,773446	0,546760	0,246323		
X13	-0,333044	-0,232967	-0,488111		
X14	0,358205	0,516527	0,071495		
X15	0,104139	-0,685161	0,073422		
X16	-0,411641	-0,216424	0,469843		
X17	0,370951	-0,626867	0,234093		
Общ. дис.	1,697535	1,862096	0,881542		
Доля обш	0,130580	0,143238	0,067811		

Рисунок 6 – Матрица факторных нагрузок

Заметим, что число сильных связей сократилось до 1, при том, что максимальное значение факторов не изменилось. Центроидный метод в факторном анализе используется для сокращения размерности данных, путем нахождения среднего значения для каждой переменной и затем использования этих средних значений для вычисления факторов. Если центроидным методом была найдена только одна единственная сильная связь фактора 1 с переменной "среднегодовая стоимость ОПФ", то можно сделать предположение, что данная переменная играет решающую роль в определении первого фактора, т.е. является наиболее значимой для объяснения вариации в данных.

Экономически интерпретировать факторы можно следующим образом:

Среднегодовая стоимость ОПФ имеет оценку 0.77 при использовании центроидного метода оценивания. Это означает, что эта переменная сильно коррелирует с единственным выделенным фактором. Таким образом, можно сделать вывод, что этот фактор отражает общие экономические условия, влияющие на переменную X_{12} . Возможно, что фактор 1 описывает эффективность использования основных средств, а среднегодовая стоимость ОПФ является одним из показателей, характеризующих затраты на основные средства. В этом случае, высокая связь между фактором 1 и среднегодовой стоимостью ОПФ может указывать на то, что компании, которые эффективно используют свои основные средства, также более экономно расходуют на них средства.

Также возможно, что среднегодовая стоимость ОПФ является показателем размера компании, а фактор 1 описывает эффективность использования ресурсов. В этом случае, высокая связь между фактором 1 и среднегодовой стоимостью ОПФ может указывать на то, что компании, которые эффективно используют свои ресурсы, могут также иметь меньшую стоимость ОПФ в среднем.

Выводы

По результатам работы можно сделать вывод о зависимости метода проведения факторного анализа и итоговых результатов исследования. Для каждого метода получена своя экономическая интерпретация факторов.

Общий итог по заданиям:

1. Методом главных компонент по исходным данным был проведен факторный анализ, не совершая поворот факторов. По корреляционной матрице можем сказать, что значения Y_2 и X_{12} имеют сильную положительную корреляцию. Также получены три общие фактора, исходя из которых, можно сделать вывод: первый фактор связан (-) с переменными «индекс снижения себестоимости продукции» и «среднегодовая стоимость ОПФ», второй фактор связан (-) с «коэффициент сменности оборудования», третий фактор связан (+) с «удельный вес рабочих в составе ППП».

Это может быть фактор, связанные с оптимизацией использования трудовых ресурсов, снижением издержек на производство и улучшением использования основных фондов.

2. На основе полученного решения произведен поворот факторов

методом Варимакс нормализованный. Также получены первые три общие фактора и дана экономическая интерпретация: первый фактор связан (-) с переменными «индекс снижения себестоимости продукции» и «среднегодовая стоимость ОПФ», второй фактор связан (-) с “коэффициент сменности оборудования”, третий фактор связан (+) с “удельный вес рабочих в составе ППП”;

3. На основе решения произведен поворот факторов методом Варимакс исходный. Получены первые три общие фактора и дана экономическая интерпретация по матрице факторных нагрузок: первый фактор связан (-) с переменными «индекс снижения себестоимости продукции» и «среднегодовая стоимость ОПФ», второй фактор связан (-) с “коэффициент сменности оборудования”, третий фактор связан (+) с “удельный вес рабочих в составе ППП”.

4. Проведен факторный анализ центроидным методом. Получен один фактора и дана экономическая интерпретация по таблице факторных нагрузок, что наиболее сильная зависимость в этом методе фактора с переменными “индекс снижения себестоимости продукции” и “ среднегодовая стоимость ОПФ ”