

УДК 330

Дата: поступления статьи / Submitted: 05.09.2019

после рецензирования / Revised: 09.10.2019

принятия статьи / Accepted: 21.11.2019



Научная статья / Scientific article

А.Ю. Трусова

Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация
E-mail: al_sun@mail.ru

Е.С. Рябова

Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация
E-mail: al_sun@mail.ru

Изучение показателей пищевой сферы Приволжского федерального округа средствами канонического анализа

Аннотация: В статье представлены результаты анализа показателей пищевой промышленности и их взаимосвязь с другими показателями социально-экономической сферы на примере Приволжского федерального округа. При анализе использован метод канонической корреляции, позволяющий многомерно оценить степень тесноты взаимосвязи. В результате выявлены группы показателей, имеющих высокую степень тесноты, а также сформированы канонические переменные.

Ключевые слова: статистический анализ, канонический анализ, коэффициенты корреляции и канонической корреляции, показатели пищевой сферы, Приволжский федеральный округ, социально-экономическая сфера.

Цитирование. Трусова А.Ю., Рябова Е.С. Изучение показателей пищевой сферы Приволжского федерального округа средствами канонического анализа // Вестник Самарского университета. Экономика и управление. 2019. Т. 10. № 4. С. 59–66.

A.Yu. Trusova

Samara National Research University, Samara, Russian Federation
E-mail: al_sun@mail.ru

E.S. Ryabova

Samara National Research University, Samara, Russian Federation
E-mail: al_sun@mail.ru

Study of indicators of food industry of the Volga Federal District by means of canonical analysis

Abstract: This article presents the results of the analysis of the food industry indicators and their relationship with different indicators of socio-economic sphere on the example of the Volga Federal District. The canonical correlation method was used in the analysis of the research, which allows one to assess the extend of strong interrelationship. Such studies as had been carried out showed that there were groups of indicators with a high level of interrelationship and canonical variables were also formed.

Key words: statistical analysis, canonical analysis, correlation coefficients and canonical correlation coefficients, food industry indicators, Volga Federal District, socio-economic sphere.

Citation. Trusova A.Yu., Ryabova E.S. Study of indicators of food industry of the Volga Federal District by means of canonical analysis. *Vestnik Samarskogo universiteta. Ekonomika i upravlenie = Vestnik of Samara University. Economics and Management*, 2019, vol. 10, no. 4, pp. 59–66. (In Russ.)

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

© Алла Юрьевна Трусова – кандидат физико-математических наук, доцент, кафедра математики и бизнес-информатики, Самарский национальный исследовательский университет имени академика С.П. Королева, 443086, Российская Федерация, г. Самара, Московское шоссе, 34.

© *Екатерина Сергеевна Рябова* – магистр I курса, специальность «Бизнес-информатика», Самарский национальный исследовательский университет имени академика С.П. Королева, 443086, Российская Федерация, г. Самара, Московское шоссе, 34.

© *Alla Yu. Trusova* – Candidate of Physical and Mathematical Sciences, associate professor, Department of Mathematics and Business Informatics, Samara National Research University, 34, Moskovskoye shosse, Samara, 443086, Russian Federation.

© *Ekaterina S. Ryabova* – Master's degree student of the 1st year of study, special field in Business Informatics, Samara National Research University, 34, Moskovskoye shosse, Samara, 443086, Russian Federation.

Введение

Основной задачей пищевой промышленности каждого государства является бесперебойное обеспечение населения качественными продуктами питания. Однако данную продукцию необходимо не только произвести, но и сбыть, следовательно, присутствует связь данного вида промышленности с такими отраслями экономики, как торговля, рынок, предпринимательство. Показатели пищевой сферы, как правило, анализируются статистическими методами с использованием монотетического подхода [1]. Но важно изучить взаимосвязь показателей пищевой промышленности с социально-экономической сферой, что возможно только благодаря комплексному подходу, который обеспечивают многомерные статистические методы.

Ход исследования

При проведении анализа использовались данные из статистического сборника, содержащего информацию о развитии экономики регионов в 2005–2016 гг. [2]. В качестве исходных данных выбрано два множества показателей X и Y (за 2010, 2013 и 2016 года) для 14 субъектов Приволжского федерального округа, которые представлены в таблице 1.

Таблица 1 – Изучаемые показатели
Table 1 – Studied indicators

Факторы X	Показатели Y
X_1 – производство молока, тыс. т	Y_1 – индексы потребительских цен на продовольственные товары, %
X_2 – производство яиц, млн шт.	Y_2 – индекс цен производителей сельскохозяйственной продукции, в % к пред. году
X_3 – производство скота и птицы, тыс. т	Y_3 – индексы приобретения товаров и услуг сельскохозяйственными организациями, в % к пред. году
X_4 – производство меда, т	Y_4 – экспорт продовольственных товаров и сельскохозяйственного сырья, млн долл. США
X_5 – валовой сбор картофеля, тыс. т	Y_5 – импорт продовольственных товаров и сельскохозяйственного сырья, млн долл. США
X_6 – валовой сбор зерна, тыс. т	Y_6 – заболеваемость на 1000 человек населения (зарегистрировано заболеваний у пациентов с диагнозом, установленным впервые в жизни)
X_7 – валовой сбор плодов и ягод, тыс. т	Y_7 – общие коэффициенты рождаемости (число родившихся на 1000 человек населения)
X_8 – валовой сбор овощей, тыс. т	Y_8 – общие коэффициенты смертности (число умерших на 1000 человек населения)
X_9 – инвестиции в с/х, охоту и лесное хозяйство, млн руб.	

В статье использовался метод анализа канонических корреляций, который позволяет одновременно анализировать взаимосвязь нескольких результативных показателей и большого числа воздействующих факторов. Анализ разбит на 3 блока [3–5].

Первый блок анализа посвящен исследованию влияния показателей пищевой сферы на экспорт и импорт продовольственных товаров и сельскохозяйственного сырья. Матрица исходных значений

переменных разбита на две части: X_5, X_6, X_7 и X_8 – переменные факторы, а Y_4 и Y_5 – результативные показатели. Расширенная матрица корреляции R для обеих групп переменных представлена в таблице 2.

Матрица R разбивается на четыре блока R_{11}, R_{22}, R_{12} и R_{21} , как показано в таблице 3.

Таблица 2 – Матрица корреляции R

Table 2 – Correlation matrix R

Фактор	X_5	X_6	X_7	X_8	Y_4	Y_5
X_5	1	0,48485	0,59207	0,669	0,23586	0,25487
X_6	0,48485	1	0,71274	0,709	0,70788	0,19349
X_7	0,59207	0,71274	1	0,836	0,79817	0,63608
X_8	0,66874	0,70905	0,83639	1	0,75559	0,59748
Y_4	0,23586	0,70788	0,79817	0,756	1	0,61201
Y_5	0,25487	0,19349	0,63608	0,597	0,61201	1

Таблица 3 – Блоки матрицы корреляции

Table 3 – Blocks of the correlation matrix

R_{11}				R_{12}	
1	0,48485	0,59207	0,669	0,23586	0,25487
0,48485	1	0,71274	0,709	0,70788	0,19349
0,59207	0,71274	1	0,836	0,79817	0,63608
0,66874	0,70905	0,83639	1	0,75559	0,59748
R_{21}				R_{22}	
0,23586	0,70788	0,79817	0,756	1	0,61201
0,25487	0,19349	0,63608	0,597	0,61201	1

К матрицам R_{11} и R_{22} находим обратные матрицы R_{11}^{-1} и R_{22}^{-1} для вычисления матрицы C , равной произведению матриц $R_{22}^{-1}, R_{21}, R_{11}^{-1}$ и R_{12} . Значения матрицы C представлены в таблице 4.

Таблица 4 – Матрица C

Table 4 – Matrix C

C	
0,807164	0,237583
0,050028	0,500705

Алгоритм канонического анализа далее предполагает вычисление собственных значений и собственных векторов для матрицы C . Размерность матрицы равна (2×2), следовательно, она имеет два собственных значения (λ_1^2 и λ_2^2) и 2 собственных вектора (B_1 и B_2).

Для вычислений в статье использовался математический калькулятор сервиса Wolfram Alpha. Таким образом, получены 2 собственных значения λ^2 , которые представлены в таблице 5, и 2 собственных вектора, соотношение:

$$B_1 = \begin{pmatrix} 6,82617 \\ 1 \end{pmatrix}, B_2 = \begin{pmatrix} -0,69615 \\ 1 \end{pmatrix}. \quad (1)$$

Квадратные корни из собственных значений матрицы C являются каноническими коэффициентами корреляции r .

Таблица 5 – Собственные значения матрицы C

Table 5 – Eigenvalues of matrix C

λ_1^2	0,841841	λ_2^2	0,4658
$\lambda_1=r_1$	0,917519	$\lambda_2=r_2$	0,6825

Компоненты векторов канонических переменных описываются векторами-столбцами A . В результате получены следующие значения векторов A :

$$A_1 = \begin{pmatrix} -4,20737 \\ 1,040815 \\ 4,653941 \\ 4,455803 \end{pmatrix}, A_2 = \begin{pmatrix} 0,103356 \\ -1,16688 \\ 0,554306 \\ 0,399374 \end{pmatrix}. \quad (2)$$

Максимальный коэффициент канонической корреляции составляет 0,92. Ему соответствуют канонические переменные:

$$U_1 = -4,21x_5 + 1,04x_6 + 0,38x_7 + 0,27x_8, \quad (3)$$

$$V_1 = 6,83y_4 + y_5. \quad (4)$$

Второму коэффициенту канонической корреляции ($r_2 = 0,68$) соответствует следующая пара канонических переменных:

$$U_2 = 0,103x_5 - 1,17x_6 + 0,55x_7 + 0,399x_8, \quad (5)$$

$$V_2 = -0,696y_4 + y_5. \quad (6)$$

Так как векторы коэффициентов канонических переменных были вычислены на основании матрицы парных коэффициентов корреляции, то они относятся к стандартизированным значениям исходных переменных [6–8].

После нормировки коэффициентов на среднее квадратическое отклонение соответствующего показателя получены две пары канонических переменных, представленных в таблице 6.

Таблица 6 – Результаты анализа первого блока показателей за 2016 год

Table 6 – Results of the analysis of the first block of indicators for 2016

R	Канонические переменные	$\chi^2_{набл.}$	$\chi^2_{кр.}$
$r_1 = 0,92$	$U_1 = -0,012x_5 + 0,0007x_6 + 0,202x_7 + 0,039x_8$	1,9833	2,73264
	$V_1 = 0,079y_4 + 0,016y_5$		
$r_2 = 0,68$	$U_2 = 0,0003x_5 - 0,0008x_6 + 0,024x_7 + 0,0035x_8$	2,84709	0,35185
	$V_2 = -0,008y_4 + 0,016y_5$		

В таблице 6 отражены результаты расчетов критерия Бартлетта при оценке значимости канонической корреляции. Как видно, на уровне значимости $\alpha = 0,05$ значимым является коэффициент канонической корреляции, равный 0,68.

Доминирующим показателем в значимой канонической корреляции при данных блоках является показатель валового сбора плодов и ягод (X_7), так как он имеет наибольший коэффициент, равный 0,024, а наименьшее влияние на показатели экспорта и импорта оказывает показатель валового сбора картофеля (X_5).

Коэффициенты при второй канонической переменной позволяют предполагать значительное снижение экспорта при еще большем увеличении импорта при изменениях в показателях пищевой сферы.

В статье представлены результаты анализа не только степени тесноты взаимосвязи показателей, но и ее изменение во времени между блоками пищевой промышленности и показателями экспорта и импорта. Далее в таблицах 7 и 8 содержатся расчеты по аналогичным группам показателей за 2010 и 2013 годы.

Таблица 7 – Результаты анализа первого блока показателей за 2010 год

Table 7 – Results of the analysis of the first block of indicators for 2010

R	Канонические переменные	$\chi^2_{\text{набл.}}$	$\chi^2_{\text{кр.}}$
$r_1 = 0,941$	$U_1 = 0,0011x_5 + 0,0024x_6 + 0,06999x_7 - 0,0133x_8$	2,11401	2,7326
	$V_1 = -0,00444y_4 + 0,01115y_5$		
$r_2 = 0,424$	$U_2 = 0,009x_5 + 0,0021x_6 + 0,1203x_7 - 0,053x_8$	1,27983	0,3519
	$V_2 = -0,08313y_4 + 0,01115y_5$		

Коэффициент r_2 при уровне значимости $\alpha = 0,05$ является статистически значимым.

Наибольшее влияние на показатели экспорта и импорта в 2010 году в значимой канонической корреляции оказывали показатели X_7 – валовой сбор плодов и ягод и X_8 – валовой сбор овощей. Наименьшее влияние оказал показатель валового сбора зерна (X_6). В 2010 году изменения в показателях пищевой промышленности могли повлечь за собой сильное снижения уровня экспорта и чуть меньшие изменения в сторону увеличения в структуре импорта.

Таблица 8 – Результаты анализа первого блока показателей за 2013 год

Table 8 – Results of the analysis of the first block of indicators for 2013

R	Канонические переменные	$\chi^2_{\text{набл.}}$	$\chi^2_{\text{кр.}}$
$r_1 = 0,907$	$U_1 = 0,01976x_5 - 0,0058x_6 + 0,01215x_7 - 0,098x_8$	2,248	2,733
	$V_1 = -0,2022y_4 + 0,00982y_5$		
$r_2 = 0,638$	$U_2 = -0,00068x_5 - 0,00127x_6 + 0,04903x_7 + 0,00208x_8$	2,632	0,352
	$V_2 = -0,00914y_4 + 0,00982y_5$		

Коэффициент $r_2 = 0,638$ является значимым на уровне значимости $\alpha = 0,05$. В 2013 году наибольшее влияние на изменение показателей экспорта и импорта оказывали валовой сбор плодов и ягод (X_7) и валовой сбор овощей (X_8). Изменения в значениях валового сбора картофеля (X_5) могли повлечь за собой наименьшие изменения в структуре экспорта и импорта. Как видно, коэффициенты при показателях экспорта (Y_4) и импорта (Y_5) приближенно равны, но различны по знаку. Это позволяет утверждать, что изменения будут одинаковыми, но различного направления.

Изучив динамику изменения значимого коэффициента корреляции, можно сделать вывод, что за изучаемый период теснота связи между показателями пищевой промышленности и показателями экспорта и импорта усилилась, так как коэффициент канонической корреляции вырос.

Для моделирования взаимосвязи показателей пищевой промышленности и показателей экономической сферы, а именно экспорта и импорта, проводилось построение регрессионных моделей на интегральных показателях (канонических переменных) за 3 временных периода, которые представлены в таблице 9.

Таблица 9 – Модели на канонических переменных
Table 9 – Models on canonical variables

2010 год	$v = 0,509u + 0,732$	$R^2 = 0,639$
2013 год	$v = 0,403u^3 - 0,5566u^2 - 0,034u - 0,0045$	$R^2 = 0,9355$
2016 год	$v = 0,802u - 0,353$	$R^2 = 0,657$

Полученные модели свидетельствуют о сильной взаимосвязи показателей и позволяют комплексно оценивать влияние блоков показателей. Данные регрессионные модели могут использоваться для визуализации данных и их прогноза на будущие периоды. Визуализация данных является важной составной частью интеллектуального анализа данных, особенно ориентированного на обработку больших объемов информации.

Было выявлено, что показатели экспорта и импорта 2013 года слабо зависят от показателей пищевой сферы 2010 года, а экспорт и импорт 2016 года имеют уже более сильную связь с показателями пищевой сферы 2013 года.

Коэффициент детерминации R^2 регрессионной модели для показателей 2010 и 2013 годов составляет 0,178, а для показателей 2013 и 2016 годов $R^2 = 0,6815$. Полученные регрессионные модели представлены формулами (7) и (8) соответственно:

$$v = 0,0208u^3 + 0,283u^2 + 1,2582u + 1,64, \tag{7}$$

$$v = 0,2905u^3 - 0,3534u^2 + 0,0584u - 0,0587. \tag{8}$$

Далее в статье проанализируем анализ взаимосвязи показателей пищевой сферы и показателей уровня здоровья населения. При значимом коэффициенте канонической корреляции (r_3) наибольшее влияние оказывают такие показатели, как производство яиц (X_2) и валовой сбор овощей (X_8), а наименьшее – показатель X_3 – производство скота и птицы на убой. Коэффициенты при показателях общих коэффициентов рождаемости и смертности очень высокие.

Это свидетельствует о том, что ухудшения в качестве и количестве производимой продукции приведут к снижению рождаемости и росту уровня смертности. Результаты расчетов представлены в таблицах 10, 11.

Таблица 10 – Канонические переменные второго блока
Table 10 – Canonical variables of the second block

Для $r_1 = 0,88$	$U_1 = -0,0051x_1 + 0,0014x_2 - 0,0072x_3 + 0,0007x_4 - 0,0021x_8 + 0,0003x_9$
	$V_1 = -0,0079y_6 - 0,196y_7 + 0,99y_8$
Для $r_2 = 0,69$	$U_2 = -0,00096x_1 + 0,0005x_2 - 0,0103x_3 + 0,0007x_4 - 0,0032x_8 + 0,00008x_9$
	$V_2 = -0,0079y_6 - 1,0309y_7 + 0,99y_8$
Для $r_3 = 0,5$	$U_3 = -0,009x_1 + 0,00348x_2 - 0,0002x_3 + 0,0012x_4 - 0,01095x_8 + 0,00065x_9$
	$V_3 = -0,032y_6 - 0,7597y_7 + 0,99y_8$

Таблица 11 – Проверка на значимость коэффициентов второго блока
Table 11 – Check on the significance of the coefficients of the second block

Показатель	Для r_1	Для r_2	Для r_3
$\chi^2_{набл.}$	1,73512	2,57099	1,30213
$\chi^2_{табл.}$	9,39046	3,9403	0,71072

Результаты анализа связи с индексами цен позволили установить, что ни один коэффициент канонической корреляции не является статистически значимым. Следовательно, нельзя утверждать, что между изучаемыми группами показателей существует тесная связь. Следовательно, пары канонических переменных не подлежат экономической интерпретации. Результаты расчетов представлены в таблицах 12, 13.

Таблица 12 – Канонические переменные третьего блока
Table 12 – Canonical variables of the third block

Для $r_1 = 0,91$	$U_1 = 0,005x_1 - 0,011x_3 + 0,0072x_5 - 0,008x_6 + 0,0002x_8$
	$V_1 = 0,3814y_1 - 0,191y_2 + 0,608y_3$
Для $r_2 = 0,42$	$U_2 = 0,0016x_1 - 0,003x_3 + 0,0029x_5 + 0,0023x_6 - 0,005x_8$
	$V_2 = -1,165y_1 + 0,1396y_2 + 0,608y_3$
Для $r_3 = 0,13$	$U_3 = -0,0066x_1 - 0,0019x_3 + 0,0157x_5 + 0,0096x_6 - 0,0102x_8$
	$V_3 = 0,633y_1 + 0,3507y_2 - 0,608y_3$

Таблица 13 – Проверка на значимость коэффициентов третьего блока
Table 13 – Checking the significance of the coefficients of the third block

Показатель	Для r_1	Для r_2	Для r_3
$\chi^2_{\text{набл.}}$	3,0375	1,2916	0,106115
$\chi^2_{\text{табл.}}$	7,260944	2,732637	0,351846

Заключение

Применение канонического анализа позволило оценить тесноту взаимосвязи показателей пищевой сферы с показателями других сфер. Выявлено, что наибольшую связь показатели пищевой промышленности имеют с показателями экспорта и импорта, так как значимый коэффициент канонической корреляции в 2016 году составил более 0,6. В результате анализа установлено, что наибольшее влияние на взаимосвязь пищевой сферы и уровня здоровья населения оказывают такие показатели, как производство молока и яиц, валовой сбор овощей. Результаты проведения канонического анализа свидетельствуют, что на уровне государственной политики актуально следить за развитием показателей пищевой сферы, так она тесно связана как с экономической, так и с социальной сферой. Таким образом, государственная политика должна быть направлена на развитие пищевой промышленности, так как ее отрасли напрямую влияют на показатели здоровья и благосостояния населения, которые в свою очередь влияют на уровень рождаемости и смертности в стране.

Библиографический список

1. Калинина В.Н., Соловьев В.И. Введение в многомерный статистический анализ: учебное пособие. Москва: ГУУ, 2003. 66 с.
2. Регионы России. Социально-экономические показатели. 2017: стат. сб. / Росстат. Москва, 2017. 1402 с. URL: https://www.gks.ru/bgd/regl/B17_14p/Main.htm.
3. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шеффер И. Многомерный статистический анализ в экономике: учеб. пособие для вузов / под ред. проф. В.Н. Тамашевича. Москва: ЮНИТИ-ДАНА, 1999. 598 с.

4. Сыровацкая И.В. Анализ канонических корреляций показателей эколого-социально-экономического развития региона // Управление экономическими системами. 2011. № 9 (33). С. 73. URL: <https://www.elibrary.ru/item.asp?id=18328600>.
5. Тихомиров Н.П., Тихомирова Т.М., Урмаев О.С. Методы эконометрики и многомерного статистического анализа. Москва: Экономика, 2017. 989 с.
6. Чураков Е.П. Введение в многомерные статистические методы: учеб. пособие. Санкт-Петербург [и др.]: Лань, 2016. 145 с.
7. Seber G.A. F. A matrix handbook for statisticians. Vol.. John Wiley & Sons, 2018, pp. 111–113. URL: http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/106_09_Seber_A_Matrix_Handbook_for_Statistic.pdf.
8. Viana A.P., Pereira M.G., Maldonado J.F.M., Junior A.T.A. Simple and canonic correlation between agronomical and fruit quality traits in yellow passion fruit // Crop Breeding and Applied Biotechnology. 2003. № 3 (2). P. 133–140. DOI: <http://dx.doi.org/10.12702/1984-7033.v03n02a06>.

References

1. Kalinina V.N., Solovyev V.I. Introduction to multivariate statistical analysis: training manual. Moscow: GUU, 2003, 66 p. (In Russ.)
2. Regions of Russia. Social and economic indicators. 2017: statistical digest. Rosstat. Moscow, 2017, 1402 p. Available at: https://www.gks.ru/bgd/regl/B17_14p/Main.htm. (In Russ.)
3. Soshnikova L.A., Tamashevich V.N., Uebe G., Shefer M. Multidimensional statistical analysis in economics: Textbook. Moscow: YuNITI-DANA, 1999, 598 p. (In Russ.)
4. Syrovatskaya I.V. Canonical correlation analysis of indicators of ecological, social and economic development of the region. *Management of Economic Systems: Scientific Electronic Journal*, 2011, no. 9 (33), p. 73. Available at: <https://www.elibrary.ru/item.asp?id=18328600>. (In Russ.)
5. Tikhomirov N.P., Tikhomirova T.M., Ushmaev O.S. Methods of econometrics and multivariate statistical analysis. Moscow: Ekonomika, 2017, 989 p. (In Russ.)
6. Churakov E.P. Introduction to multidimensional statistical methods: textbook. Saint Petersburg [et al.]: Lan', 2016, 145 p. (In Russ.)
7. Seber G.A.F. A matrix handbook for statisticians. Vol. 15. John Wiley & Sons, 2018, p. 111–113. Available at: http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/106_09_Seber_A_Matrix_Handbook_for_Statistic.pdf.
8. Viana A.P., Pereira M.G., Maldonado J.F.M., Junior A.T.A. Simple and canonic correlation between agronomical and fruit quality traits in yellow passion fruit. *Crop Breeding and Applied Biotechnology*, 2003, vol. 3, no. 2, pp. 133–140. DOI: <http://dx.doi.org/10.12702/1984-7033.v03n02a06>.