

УДК 332.012

***АНАЛИЗ ВЗАИМОСВЯЗИ ВАЛОВОГО РЕГИОНАЛЬНОГО
ПРОДУКТА И ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ РЕГИОНА МЕТОДОМ
МАТЕМАТИЧЕСКОЙ РЕГУЛЯЦИИ***

Кочкин Т.Н.,

бакалавр

Санкт-Петербургский государственный экономический университет

Санкт-Петербург, Россия

Аннотация. В статье произведено исследование взаимосвязи макроэкономических и социальных показателей и их влияние на значение валового регионального продукта. В рамках нашей научной работы ВРП был выбран нами как основной показатель, характеризующий благополучие субъекта РФ. Целью исследования является выделение показателей, которые оказывают наибольшее влияние. Помимо основных показателей, которые выделяют в подобных работах, как показатель среднедушевого дохода населения, мы решили также использовать такие показатели, которые обычно остаются без внимания, как коэффициент среднего количества населения в браке в регионе. Основными инструментами являются методы математической статистики, линейной регрессии, методами математической регуляризации и машинного обучения.

Ключевые слова: Регионы Российской Федерации; валовый региональный продукт; статистический анализ данных; прогноз показателя; качество жизни населения; региональная политика; социально-экономическое положение регионов.

***ANALYSIS OF THE RELATIONSHIP OF THE GROSS REGIONAL
PRODUCT AND THE ECONOMIC INDICATORS OF THE REGION BY THE
METHOD OF MATHEMATICAL REGULATION***

Kochkin T.N.,

Bachelor

Saint-Petersburg State University of Economics

Saint Petersburg, Russia

Annotation. The article studies the relationship of macroeconomic and social indicators and their impact on the value of gross regional product. In the framework of our scientific work, GRP was chosen by us as the main indicator characterizing the well-being of the subject of the Russian Federation. The aim of the study is to highlight the indicators that have the greatest impact. In addition to the main indicators that are distinguished in such works as an indicator of the average per capita income of the population, we also decided to use indicators that are usually ignored as the coefficient of the average number of people married in the region. The main tools are methods of mathematical statistics, linear regression, methods of mathematical regularization and machine learning.

Keywords: Regions of the Russian Federation; gross regional product; statistical data analysis; indicator forecast; quality of life of the population; regional policy; socio-economic situation of the regions.

Целью исследования является определение основных экономических и социальных показателей, которые оказывают наибольшее влияние на ВРП региона, а также разработка модели, которая позволит рассчитать ВРП на основании известных показателей региона [7]. Для нашего исследования мы проанализировали массив статистических данных по всем субъектам Российской Федерации [5]. Актуальность темы можно определить следующим образом: большинство исследований не пользуются современными методами математической статистики и машинным обучением [1]. Мы же применяем актуальные математические методы, которые позволяют получать точные

Вектор экономики | www.vectoreconomy.ru | СМЭЛ № ФС 77-66790, ISSN 2500-3666

данные и делать верные выводы из полученных результатов [8]. На основании результатов государственные органы смогут точнее регулировать региональную политику, что позволит увеличить значимые макроэкономические показатели и улучшить качество жизни в регионах.

Валовый региональный продукт является одним из основных показателей, характеризующих экономическое благополучие региона [8]. Высокий показатель ВРП позволяет судить о высоком уровне жизни населения, что характеризуется высокими среднедушевыми доходами, потреблением платных услуг в регионе и т. д. При помощи математических методов мы собираемся проверить данные взаимосвязи, чтобы гарантировать их существование и точность [2]. Государственные органы при разработке политики и планировании регионального бюджета зачастую полагаются на статистические данные и выводы, сделанные по ним, поэтому необходима методика, позволяющая определить основные показатели, на которые необходимо обращать внимание в первую очередь [9]. Для нахождения взаимосвязи между различными социально-экономическими показателями регионов и показателем ВРП региона мы будем использовать коэффициент корреляции Пирсона.

В качестве базовой модели взаимосвязи ВРП региона и остальных социально-экономических факторов выберем модель линейной регрессии:

$$y = \sum_{i=1}^n w_i x_i + w_0 + \varepsilon \quad (2)$$

Где x_i – независимые социально-экономические факторы модели

w_i – весовые коэффициенты факторов модели

y – зависимая переменная, соответствующая ВРП региона

w_0 – свободный член

ε – неустранимая ошибка

n – количество независимых факторов в модели

Согласно методологии обучения с учителем мы имеем набор из l объектов $(x_1^1 \dots x_n^1), \dots (x_1^l \dots x_n^l)$, для которых дан набор известных значений зависимой переменной y^1, \dots, y^l . На основании этих данных нужно восстановить зависимость (2). В данной работе мы будем активно пользоваться матрично-векторной формой записи для более компактного представления получаемых формул. В матричной форме модель (2) выглядит следующим образом:

$$y = Xw + \varepsilon \quad (3)$$

Где $y = (y^1, \dots, y^l)^T$

$$X = \begin{pmatrix} 1 & \dots & x_n^1 \\ \dots & \dots & \dots \\ 1 & \dots & x_n^l \end{pmatrix}$$

$$w = (w_0, \dots, w_n)^T$$

ε – неустраняемая ошибка

Весовые коэффициенты w_i определяются согласно методу наименьших квадратов:

$$Q(w, X) = \frac{1}{2l} \sum_{i=1}^l (y^i - w^T X_i)^2 = \frac{1}{2l} (y - Xw)^T (y - Xw) \xrightarrow{w} \min \quad (4)$$

Где X_i – i -тая строка матрицы X .

Условие минимума данного функционала (4):

$$\frac{\partial Q}{\partial w} = 0 \quad (5)$$

Из условия (5) можно получить формулу для вычисления искомым весовых коэффициентов:

$$w = (X^T X)^{-1} X^T y \quad (6)$$

Имеющийся у нас набор данных мы разбили на обучающую и тестовую выборки. Модель будет строиться на основании обучающей выборки, а качество полученной модели (2) будет оцениваться с помощью коэффициента детерминации как на обучающей выборке, так и на тестовой. Величина коэффициента детерминации соответствует величине дисперсии зависимой переменной, которая объясняется моделью (2).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y^i - w^T X_i)^2}{\sum_{i=1}^n (y^i - \bar{y})^2} \quad (7)$$

Учитывая разный масштаб имеющихся социально-экономических показателей – значения некоторых меньше 1, а других больше 1000, для построения качественной модели машинного обучения необходимо произвести масштабирование имеющихся данных.

Масштабирование будет применяться ко всем независимым факторам по следующей формуле:

$$x_i^j \rightarrow \frac{x_i^j - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (8)$$

Где x_i^j – значение i -того фактора j -того объекта

$\max(x_i)$ – наибольшее значение i -того фактора в выборке

$\min(x_i)$ – наименьшее значение i -того фактора в выборке

Данное преобразование гарантирует принадлежность всех факторов всех объектов отрезку $[0;1]$.

Для отбора значимых показателей попробуем ограничить величину весов в модели. Для этого мы добавим в функционал (4) регуляризующее слагаемое, начав с L2 регуляризации:

$$Q(w, X) = \frac{1}{2l} (y - Xw)^T (y - Xw) + \alpha \|w\|_2^2 \rightarrow \min_w \quad (9)$$

Где α – вещественный параметр, влияющий на величину регуляризации

$$\|w\|_2 - \text{L2 векторная норма, т.е. } \|w\|_2 = \sqrt{\sum_{i=1}^n w_i^2}$$

В нашей модели мы рассматривали 11 независимых социально-экономических показателей, которые приведены в таблице 1. По формуле (1) были рассчитаны коэффициенты корреляции Пирсона между каждым из 11 показателей и показателем ВРП региона. В таблице 1 приведены результаты расчёта.

Таблица 1. Результаты расчёта коэффициента корреляции Пирсона.

Показатель	r_{xy}
Население трудоспособного возраста	0.20
Население, состоящее в браке	-0.05
Работающее население	0.54
Среднедушевой доход населения	0.82
Инвестиции на душу населения	0.88
Уровень преступности	0.17
Объём платных услуг на душу населения	0.74
Удельный вес городского населения	0.43
Население	0.26
Промышленная добыча на душу населения	0.85
Промышленная обработка на душу населения	0.28

Источник: Рассчитано автором на основании статистических данных по регионам Российской Федерации.

На обучающей выборке коэффициент детерминации (7) равняется 0.97, тогда как на тестовой 0.95. Это говорит о высоком качестве модели и об отсутствии такого явления, как переобучение модели. Значит построенная модель (2) в целом адекватно представляет взаимосвязь группы социально-

экономических показателей региона и ВРП региона [3]. Для отбора наиболее связанных с ВРП социально-экономических показателей регионов применим описанные ранее методы регуляризации задачи (4). Мы перебрали несколько значений параметра α для модели с L2-регуляризацией [10]. Ниже приведена таблица значений коэффициентов детерминации моделей с различными параметрами α .

Таблица 2. Качество моделей с L2-регуляризацией.

α	R^2 на обучающей выборке	R^2 на тестовой выборке
1.0	0.97	0.95
5.0	0.93	0.92
10.0	0.89	0.86

Источник: Рассчитано автором на основании построенных моделей.

Из данной таблицы видно, что с ростом параметра регуляризации α качество модели ухудшается как на обучающей выборке, так и на тестовой.

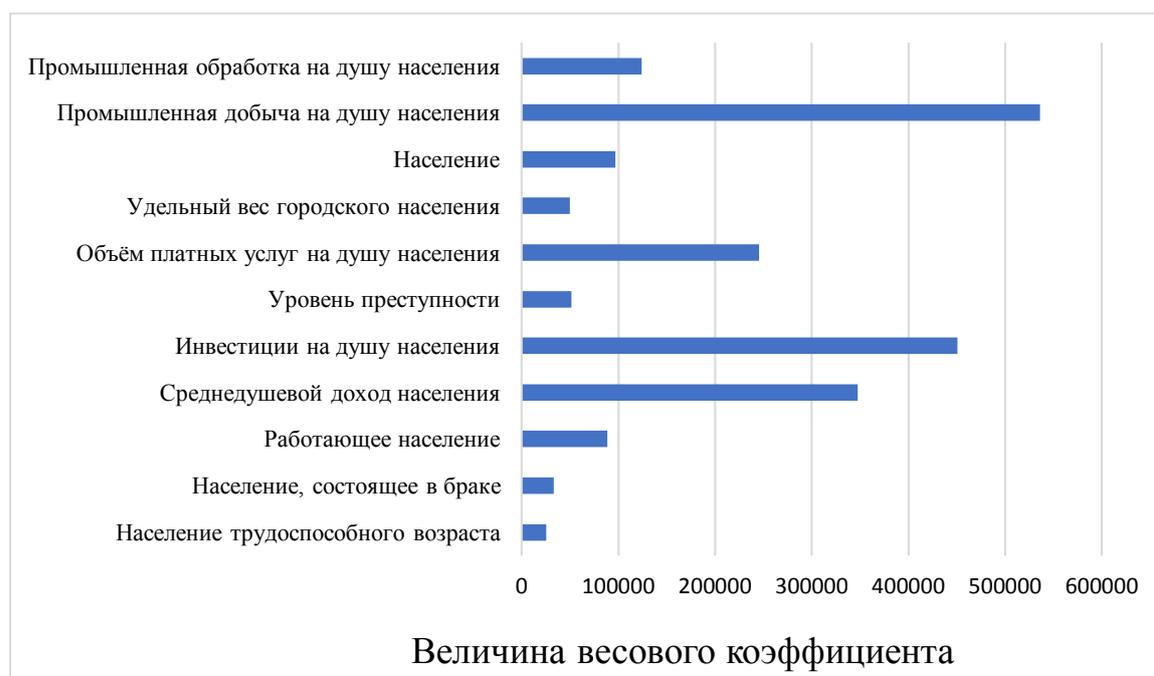


Рис. 1. Весовые коэффициенты регуляризированной модели.

На рисунке 1 видно, что произошло уменьшение весовых коэффициентов у некоторых показателей, однако мы не получили нулевых коэффициентов, поэтому в данной задаче L2-регуляризация не приводит ни к улучшению модели,

ни к отбору наиболее значимых для ВРП социально-экономических показателей. Поэтому мы используем L1-регуляризацию [4].

Мы перебрали различные значения параметра α из отрезка $[0.1, 1000]$, и для всех из них качество модели почти не меняется и равно:

R^2 на обучающей выборке – 0.96

R^2 на тестовой выборке – 0.95

Такие значения коэффициента детерминации позволяют говорить о хорошем качестве регуляризированной модели.

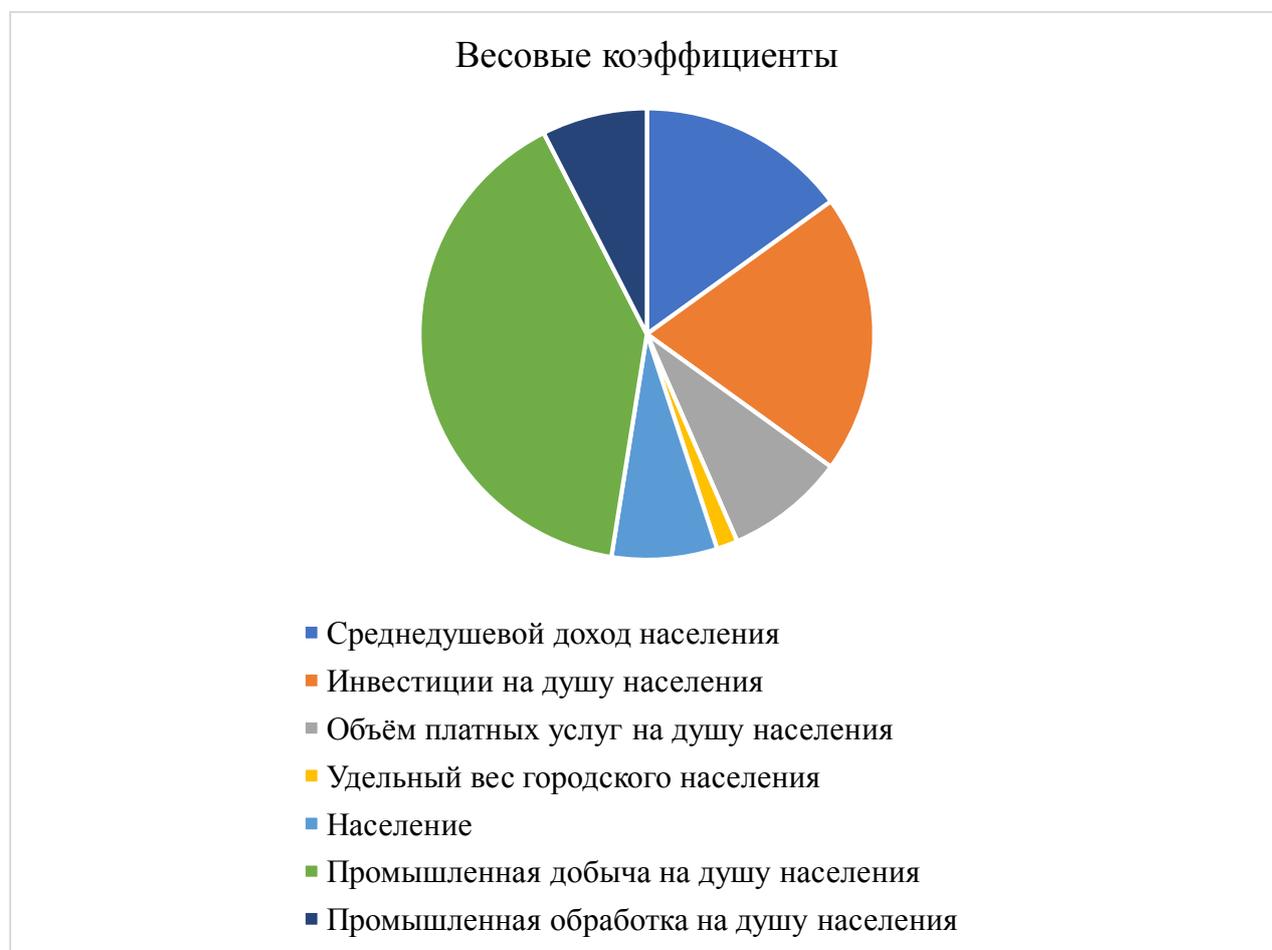


Рис. 2. Весовые коэффициенты регуляризированной модели.

На рисунке 2 приведено сравнение весовых коэффициентов значимых для прогноза ВРП. Такие показатели как население трудоспособного возраста,

население, состоящее в браке, работающее население, уровень преступности – не оказывают значимого влияния на прогноз ВРП [11]. Таким образом применение L1-регуляризации в данной задаче позволило не только построить модель с высоким качеством на тренировочной и тестовой выборках, но и отобрать наиболее значимые социально-экономические показатели регионов для прогноза ВРП региона [12].

Проведенное нами исследование позволило нам выделить основные показатели, влияющие на значение валового регионального продукта, а именно: промышленная добыча, среднедушевой доход, инвестиции и объем платных услуг на душу населения, при помощи весовых коэффициентов мы можем оценить вклад каждого показателя в итоговый – ВРП [13]. Но несмотря на то, что эти показатели имеют наибольший удельный вес и требуют наибольшего внимания от региональных властей, им также не стоит забывать об улучшении других экономических и социальных показателей региона [5]. Стоит отметить, что наибольший вклад в величину ВРП вносит промышленная добыча, что подтверждает сильную зависимость экономики России от добычи углеводородов, газа нефти и прочих полезных ископаемых, и слабое развитие высокотехнологичной промышленности и сферы услуг [14]. Государству необходимо корректировать текущую экономическую политику, чтобы обеспечить улучшение экономических и социальных показателей, чтобы обеспечить населению высокий уровень жизни.

Библиографический список:

1. Баширова А.А. К вопросу о понятии эколого-экономической сбалансированности развития региона / А.А. Баширова // Апробация. – 2015. – № 8 (35). – С. 50-52.
2. Блищак К.А. Развитие несырьевого сектора экономики в России как основа сбалансированности бюджета / К.А. Блищак, М.Л. Вилисова // Синергия Наук. – 2018. – № 19. – С. 186-190.

3. Болтов Р.В. Сбалансированность развития территорий / Р.В. Болтов, О.И. Карелин, А.В. Семашко // Евразийский союз ученых. – 2016. – No 4 (25). – С. 109-110.
4. Борисова И. С. Проблемы устойчивого развития экономики региона в условиях преобладания отдельного вида хозяйственной деятельности / И. С. Борисова. – М.: Буки Веди, 2016. – 288 с.
5. Borkova E., Rasputin A., Vostrikov N., Kochkin T., Kartavchenko A. Analysis of the Relationship of the Volume of Paid Services for Capita and Socio-Economic Indicators of the Region (Case of Russia) // 34th IBIMA Conference: 13-14 November 2019, Madrid, Spain
6. Боркова Е.А., Городкова С.А., Ватлина Л.В., Никифоров А.А. Мониторинг экономического состояния региона на пути инновационного развития России. Монография в 2 частях / Чита, 2015. Том Часть 2. – 200 с.
7. Бородушко Е.С. Проблемы сбалансированности и безопасности социально-экономического развития регионов России / Е.С. Бородушко // Вестник Санкт-Петербургской юридической академии. – 2012. – No 1 (14). – С. 81-86.
8. Дробышевская Л.Н. Оценка структурной сбалансированности инвестиционного развития регионов / Л.Н. Доброшевская, К.И. Шевченко // Региональная экономика: теория и практика. – 2014. – No 8 (335). – С. 34-43.
9. Колосовский Н.Н. Основы экономического районирования / Н.Н. Колосовский. – М.: Госполитиздат, 1958. – 200 с.
10. Красиков Н.В. Сбалансированность развития отраслей в региональной экономике (на примере архангельской области) / Н.В. Красиков // Управление экономическими системами: электронный научный журнал. – 2012. – № 9 (45). – С. 50.
11. Миненко А.В. Проблемы моделирования инфраструктуры развития малого предпринимательства сельского муниципального образования / А.В. Миненко // Вектор экономики. - 2018. - №10.

12. Нестеров А.Н. Моногород: управление развитием / А.Н. Нестеров, Л.Г. Иогман, Н.Ю. Литвинова, С.Н. Ткачук, Т.В. Ускова. – Вологда: ИСЭРТ РАН, 2012. – 220 с.

13. Селищева Т.А. Причины затухающей динамики экономики России // В сборнике: Социально-экономическое развитие России и Монголии: проблемы и перспективы Материалы VI Международной научно-практической конференции. 2019. С. 296-300.

14. Селищева Т.А. Проблемы устойчивого развития трансграничных регионов Сибири и дальнего востока в евразийской интеграции // Проблемы современной экономики. - 2019. - № 1 (69).- С. 9-15.

Оригинальность 80%