

ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

*Д.И. Пионтковский, Д.Д. Соколов,
О.С. Старчикова*

**СРАВНЕНИЕ МАТЕМАТИЧЕСКИХ
МЕТОДОВ ПРОГНОЗИРОВАНИЯ
ТАБЛИЦ «ЗАТРАТЫ – ВЫПУСК»
НА ОСНОВЕ БАЗЫ ДАННЫХ WIOD**

Препринт WP2/2015/07

Серия WP2

Количественный анализ в экономике

Москва
2015

Редактор серии WP2
«Количественный анализ в экономике»
В.А. Бессонов

Пионтковский, Д. И., Соколов, Д. Д., Старчикова, О. С.

Сравнение математических методов прогнозирования таблиц «затраты – выпуск» на основе базы данных WIOD [Электронный ресурс] : препринт WP2/2015/07/ Д. И. Пионтковский, Д. Д. Соколов, О. С. Старчикова ; Нац. исслед. ун-т «Высшая школа экономики». – Электрон. текст. дан. (1 МБ). – М. : Изд. дом Высшей школы экономики, 2015. – (Серия WP2 «Количественный анализ в экономике»). – 37 с.

Исследуется сравнительная эффективность методов проекции таблиц «затраты – выпуск» применительно к таблицам использования. Эмпирической базой исследования служили таблицы использования 28 стран мира за 1995–2010 гг. по данным международного проекта WIOD. Проведено сравнительное исследование трех математических методов, показавших себя наиболее эффективными при построении проекций таблиц использования Испании и Нидерландов в эмпирическом исследовании Темуршоева, Вебба и Яmano (2011). Наиболее эффективным из этих методов в нашем исследовании оказался бипропорциональный метод GRAS, результаты применения которого по ряду критериев оказались более близкими к опубликованным таблицам, чем результаты применения метода INSD и метода Куроды, основанных на квадратичном программировании. Это позволяет рассматривать метод GRAS как приоритетный при экстраполяции таблиц использования в экономике России.

Ключевые слова: проект WIOD, RAS, метод Куроды, таблицы «затраты – выпуск»

Классификация JEL: D57, C53, C67

Статья подготовлена в результате проведения исследования/работы (№ 15-05-0020) в рамках Программы «Научный фонд Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ)» в 2014-2015 гг. и с использованием средств субсидии на государственную поддержку ведущих университетов Российской Федерации в целях повышения их конкурентоспособности среди ведущих мировых научно-образовательных центров, выделенной НИУ ВШЭ

**Препринты Национального исследовательского университета
«Высшая школа экономики» размещаются по адресу: <http://www.hse.ru/org/hse/wp>**

© Пионтковский Д. И., 2015
© Соколов Д. Д., 2015
© Старчикова О. С., 2015
© Оформление. Издательский дом
Высшей школы экономики, 2015

1. Введение

Таблицы «затраты – выпуск», впервые предложенные в 1930-е годы В.В. Леонтьевым, используются во множестве прикладных исследований и являются частью системы национальных счетов (СНС) многих стран мира [Miller, Blair, 2009].

Идея этих таблиц состоит в том, чтобы отразить в одной таблице как можно более полную картину производства различных продуктов предприятиями всех отраслей экономики, а также потребления всеми отраслями всех производимых ими же продуктов внутри одной экономики в течение года. Первая таблица носит название таблицы ресурсов, а вторая – таблицы использования товаров и услуг: именно с последними мы будем иметь дело в настоящей работе.

Таблицы использования строятся либо в основных (базовых) ценах, либо в ценах покупателей. Таблица использования в ценах покупателей вычисляется как сумма следующих таблиц: использования отечественных товаров и услуг в основных ценах, использования импортных товаров и услуг в основных ценах (сумма этих двух – таблица использования в основных ценах), транспортных наценок, торговых наценок и чистых налогов на продукты. В этой работе будут проведены сравнения методов проекции на основе таблиц и в ценах покупателей, и в основных ценах.

Таблица использования разбивается на четыре квадранта, устроенных следующим образом: в квадранте I отражено потребление каждой задействованной в экономике отрасли всех видов производимых продуктов (промежуточное потребление); в квадранте II представлен конечный спрос в разных его элементах, а в квадранте III – добавленная стоимость каждой отрасли. Квадрант IV чаще всего не заполняется, но иногда в нем (частично) отражается перераспределение национального дохода. Таблицы «затраты – выпуск» строятся статистическими службами отдельных стран: такие таблицы могут объединяться в таблицы для групп стран (например, для Евросоюза).

Ниже представлена таблица использования товаров и услуг [Miller, Blair, 2009].

Таблица 1. Использование товаров и услуг

	Отрасли	
Продукты	Квадрант I	Квадрант II
	Промежуточное использование: X_{ij} – количество продукта i , необходимое для производства единицы продукции отрасли j ($X_{ij} \geq 0 \forall i, j$).	Конечный спрос: потребление домашних хозяйств, потребление государства и некоммерческих организаций, обслуживающих домашние хозяйства, валовое накопление основного капитала, изменение запасов, чистый экспорт (экспорт минус импорт)
	Квадрант III	Квадрант IV
	Добавленная стоимость (оплата труда, налоги на производство за вычетом субсидий, предпринимательский доход, валовая прибыль)	Элементы перераспределения национального дохода

Построение системы таблиц «затраты – выпуск» требует масштабного обследования предприятий всех значимых отраслей, которое обычно проводится национальными статистическими службами не чаще, чем один раз в несколько лет (как правило – один раз в пять лет). Поскольку прикладные задачи анализа и прогноза требуют временного ряда ежегодных таблиц, то на основе таблиц, построенных в результате упомянутого обследования (так называемых базовых таблиц), и ежегодных данных национальных счетов строятся таблицы за годы, когда такого обследования не проводилось. Эта процедура носит название проекции таблиц.

Для этой процедуры – проекции таблиц «затраты – выпуск» – разработан ряд математических методов, среди которых наиболее известным является классический метод RAS (см. о нем в разделе 2.2 ниже). Эти методы постоянно совершенствуются, разрабатываются новые, и для выявления наиболее эффективных методов регулярно проводятся сравнительные эмпирические исследования (краткий обзор работ в этой области см. в [Miller, Blair, 2009, 7.4]). Вопрос о проекции таблиц «затраты – выпуск» актуален сейчас для России, поскольку на основе базовой системы таблиц за 2011 г., которую Росстат предполагает выпустить до конца 2015 г., предстоит построить перспективный и ретроспективный ряд аналогичных таблиц за последующие годы (до года, за который будет выпущена следующая базовая таблица) и за предшествующие годы (до 1995 г., за который имеется предыдущая базовая таблица). При построении таблиц за 2003–2010 гг. по России

наиболее интересной представляется задача построения системы таблиц использования [Баранов и др., 2014].

Одно из последних сравнительных эмпирических исследований методов проекции таблиц «затраты – выпуск» предпринято в [Temurshoev, Webb, Yamano, 2011]. В нем результаты проекций по 10 методам сравнивались с реальными данными на таблицах различного типа по Испании и Нидерландам. По результатам этого исследования наиболее эффективной показала себя соответствующая модификация метода RAS – так называемый метод GRAS (см. ниже). Вместе с тем именно на интересующих нас таблицах использования результаты не выглядят однозначными. При проекции таблиц использования не менее эффективными, чем методы RAS и GRAS, показали себя два квадратичных метода (они подробно описаны ниже) – метод INSD и метод Куроды. Так, на данных Нидерландов при построении таблицы использования в ценах покупателей 2000 г. по данным 1995 г. наиболее эффективным оказался метод Kuroda 1, а при проекции таблицы 2000 г. на 2005 г. – метод RAS [Temurshoev, Webb, Yamano, 2011, Table 3]. Аналогично при проекции таблиц использования Испании в основных ценах с 2000 г. на 2005 г. в первом квадранте лучше оказался метод Куроды, а во втором квадранте – метод RAS (см. два последних графика на Fig.1 и Table 4 в *Op. cit.*).

Цель настоящей статьи – более полное сравнительное эмпирическое исследование методов проекции таблиц использования. Отправной точкой мы считаем описанные результаты Темуршоева, Вебба и Яmano: так, мы полагаем установленным, что наиболее эффективными на сегодняшний день можно считать методы GRAS, INSD и один из вариантов метода Куроды, и исследуем только эти методы. Поскольку, к тому же, мы концентрируемся на таблицах использования, все это позволило расширить эмпирическую базу исследования: мы использовали данные международного проекта WIOD, в рамках которого собраны таблицы использования по 40 странам мира за 1995–2012 гг. (см. www.wiod.org). Вместе с тем нам приходится учитывать, что при построении таблиц использования на основе базовых таблиц национальные статистические службы также используют математические методы (см., например, [Eurostat, 2008, 8.6.1]). Тем самым простое сравнение методов прогнозирования на таком эмпирическом материале может привести к искажениям: например, если официальная таблица за 2001 г. построена на основе таблицы за 2000 г. методом RAS, то, естественно, метод RAS покажется наиболее эффективным при проекции с 2000 г. на 2001 г., но никакой информации о его релевантности применительно к прогнозированию реального состояния экономики это не даст. Аналогичным образом если за какие-то годы официальные таблицы «затраты – выпуск» отсутствовали, то недостающие таблицы строились в рамках проекта WIOD с помощью

проекции [Erumban et al., 2012], в частности, посредством предложенного в [Temurshoev, Timmer, 2011] варианта метода RAS. Например, все таблицы по России построены этим методом на основе таблиц за 1995 г., поэтому, как и показали наши вычислительные эксперименты, метод RAS оказывается при проекции таблиц использования по России (с 1995 на 2000-е годы) более точным, чем на примере других стран.

Чтобы избежать подобных статистических артефактов, было принято решение использовать только те таблицы из базы WIOD, которые построены путем агрегации базовых национальных таблиц. Если при этом источники данных проекта WIOD исчерпывающе описаны в [Erumban et al., 2012], то, к сожалению, достоверных сведений о том, какие именно из построенных национальными статистическими службами таблицы были базовыми, по некоторым странам найти не удалось. За исключением этих стран, а также тех, в которых, подобно России, за рассматриваемый период была построена только одна базовая таблица, осталось 28 стран, данные по которым и составили эмпирическую базу нашего исследования.

В результате исследования оказалось, что наиболее эффективным из всех трех методов оказался традиционный метод GRAS – обобщение классического метода RAS на матрицы с возможными отрицательными элементами.

Структура работы следующая. Во втором разделе мы напоминаем описание трех рассматриваемых здесь методов проекции таблиц «затраты – выпуск», а также нескольких традиционных «метрик», позволяющих сравнивать между собой прогнозы и прогнозы с реальными данными. В основном мы следуем здесь статье [Temurshoev, Webb, Yamano, 2011]. В третьем разделе работы описаны исходные данные нашего исследования. Существенную часть здесь составляет информация о том, за какие годы включенные в WIOD страны выпускали базовые таблицы использования. На основе этой информации мы выбираем таблицы использования по 28 странам из базы данных WIOD для вычислительных экспериментов, результаты которых представлены в разделе 4. Заключение посвящено краткому анализу этих результатов.

Благодарности

Мы благодарны коллегам по учебно-научной группе «Затраты – выпуск» НИУ ВШЭ, прежде всего Э.Ф. Баранову, за постановку задачи и ценные замечания, Е.А. Старицыной и Е. Корневой за постоянное внимание к нашей работе, и, в особенности С. Кузнецову и М. Родичкину, с которыми мы обменивались опытом по реализации математических методов проекции таблиц. Мы также благодарны коллегам, которые прояснили нам ситуацию с базовыми годами составления таблиц использования в различных странах, в том

числе *Mehran Kafai* (Statistics Portal Grand Duchy of Luxembourg), *Eva Schwarz* (Federal Statistical Office Germany) и *Ylva Petersson Strid* (Statistics Sweden).

2. Методы проекции и сопоставление результатов их применения

2.1. Общая методология проекции таблиц и сравнение результатов прогнозов

Мы будем рассматривать проекцию следующих возможных частей таблицы использования: квадранта I таблицы (промежуточное потребление), квадранта II (конечный спрос) или квадрантов I и II вместе.

Предположим, что у нас имеется матрица A_g^s – соответствующая часть таблицы использования, построенной для страны s за год g , исходя из статистических данных. Также предположим, что известны (из национальных счетов или из других источников) окаймляющие итоги строк и столбцов аналогичной матрицы A_{g+t}^s (которая полностью состоит из неизвестных) за год $(g + t)$, где $t = 1, 2, \dots$. Пусть размерность прогнозируемой матрицы равна $n \times m$ (в формате таблиц WIOD, которые мы рассматриваем в этой статье, $n = 59$, а $m = 36, 6$ или 42 в зависимости от того, рассматриваем ли мы первый квадрант, второй квадрант или же оба вместе). Теперь построим матричное представление системы описанных выше линейных ограничений на матрицу неизвестных A_{g+t}^s

$$(1) \quad G \cdot \vec{a}_{g+t}^s = \vec{c}_{g+t}^s,$$

где G – это матрица размера $(n + m) \times (n \cdot m)$, состоящая из нулей и единиц; \vec{a}_{g+t}^s – векторизация матрицы неизвестных A_{g+t}^s (размера $(n \cdot m) \times 1$); \vec{c}_{g+t}^s – вектор размеров $(n + m) \times 1$, составленный из векторов окаймляющих итогов строк и столбцов матрицы A_{g+t}^s .

Искомая матрица должна удовлетворять системе (1). Но для поиска решения нам необходимо начальное приближение, которое было бы сходно с искомым по структуре. Таким приближением выступит векторизация известной нам матрицы A_g^s . Обозначим ее \vec{a}_g^s .

Наконец, мы передаем посчитанные значения G, \vec{a}_g^s и \vec{c}_{g+t}^s в процедуру прогнозирования и на выходе получаем прогноз \vec{A}_{g+t}^s матрицы использования.

Теперь перейдем к сравнению результатов прогнозирования двух произвольных методов M_1 и M_2 .

Нужно сразу оговориться, что для корректного сравнения нам понадобится не только знание таблицы использования для начального года g , но и знание такой таблицы для целевого года $(g + t)$, поскольку мы будем в конкретном вычислительном эксперименте

рассматривать как более эффективный тот метод прогнозирования, результат применения которого ближе, в некотором смысле, к истинной матрице A_{g+t}^S .

Упомянутая близость двух матриц может быть формализована различными способами. Здесь мы, следуя [Temurshoev, Webb, Yamano, 2011], рассматриваем семь следующих метрик. Все эти метрики хорошо известны (см., в частности, [Miller, Blair, 2009, 7.4.8]). Мы следуем обозначениям из [Temurshoev, Webb, Yamano, 2011, section 3]: в каждом из семи случаев вычисляется своя «метрика» между матрицами (x_{ij}) и (x_{ij}^{true}) . В дальнейшем мы вычисляем их для случая, когда $(x_{ij}^{true}) = A_{g+t}$ – истинная матрица прогнозируемого года, а $(x_{ij}) = A_{(g+t),m}$ – матрица прогноза на нужный нам год, где m – номер метода.

- 1) **Средняя абсолютная ошибка в процентах**
(MAPE – Mean absolute percentage error):

$$MAPE = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m \frac{|x_{ij} - x_{ij}^{true}|}{|x_{ij}^{true}|} \times 100$$

- 2) **Взвешенная абсолютная процентная ошибка**
(WAPE – Weighted absolute percentage):

$$WAPE = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{|x_{ij}^{true}|}{\sum_k \sum_l x_{kl}^{true}} \right) \frac{|x_{ij} - x_{ij}^{true}|}{|x_{ij}^{true}|} \times 100$$

- 3) **Стандартизированная взвешенная абсолютная разница**
(SWAD – Standardized weighted absolute difference):

$$SWAD = \frac{\sum_{i=1}^n \sum_{j=1}^m |x_{ij}^{true}| \times |x_{ij} - x_{ij}^{true}|}{\sum_k \sum_l (x_{kl}^{true})^2}$$

- 4) **Пси-статистика (PsiStat):**

$$\hat{\psi} = \frac{1}{\sum_k \sum_l x_{kl}^{true}} \sum_i \sum_j \left[|x_{ij}^{true}| \times \left| \ln \left(\frac{x_{ij}^{true}}{s_{ij}} \right) \right| + |x_{ij}| \times \left| \ln \left(\frac{x_{ij}}{s_{ij}} \right) \right| \right],$$

где $s_{ij} = (|x_{ij}^{true}| + |x_{ij}|)/2$.

- 5) **RSQ** – коэффициент корреляции между построенной матрицей и целевой матрицей.

- 6) **Невязка** ($I_{\text{нас}}$) – максимум модуля разности между окаймляющими итогами построенной и целевой матриц.
- 7) **$N0$** – количество нулевых элементов матрицы прогноза, которые соответствуют ненулевым элементам в истинной матрице [Temurshoev, Webb, Yamano, 2011, section 3]. В наших вычислительных экспериментах оказалось, что значение $N0$ практически всегда нулевое, так что в дальнейших расчетах оно не учитывалось.

В результате мы получаем для каждой пары матриц A_g и A_{g+t} по три значения каждой из метрик (по одной на метод). Теперь проранжируем эти три значения для всех метрик, кроме **$N0$** (потому что она будет одинаковой): для метрик 1), 2), 3), 4) и 6) – по убыванию (чем меньше – тем лучше), а для метрики 5) – по возрастанию (так как корреляция увеличивается при увеличении сходства). В итоге мы получаем три шестимерных вектора рангов R_m для каждого из трех методов ($m = 1, 2, 3$): i -я компонента вектора R_m равна рангу метода m среди трех методов относительно метрики i , где ранг 1 соответствует наиболее эффективному методу, а ранг 3 – наименее эффективному.

Наконец, вычислим теперь суммы элементов векторов R_1, R_2 и R_3 и обозначим их через r_1, r_2 и r_3 соответственно. Сравним эти три числа, ранжируем три метода и определяем наиболее эффективный из них применительно к проекции матрицы A_g на год $g+t$: более эффективному методу соответствует наименьшее значение r_i , а если r_i и r_j оказались равны, то считаем, что оба метода одинаково эффективны.

2.2. Некоторые методы проекции таблиц «затраты – выпуск»

Дадим формальное описание исследуемых методов проекции.

I. Метод *GRAS*

Метод *GRAS*, изначально предложенный в [Gunluk-Senesen, Bates, 1988], является прямым потомком классического метода *RAS* Ричарда Стоуна [Stone, 1961]. Для начала определим информацию, которая задается экзогенно перед началом построения прогноза. Очевидно, что прежде всего понадобится матрица «затрат-выпуска» за базовый год g , затем нам нужна будет какая-нибудь информация, вычисленная для прогнозируемого года ($g + t$). Этой информацией будут окаймляющие итоги квадрантов таблицы за год ($g + t$), которую еще предстоит построить. Как было сказано выше, эти окаймляющие итоги можно получить из статистических данных (национальных счетов).

Теперь дадим описание самих методов. При выполнении методов *RAS* и *GRAS* [Temurshoev, Miller, Bouwmeester, 2013] строятся два вектора множителей: вектор-столбец \vec{r} и вектор-строка \vec{s} , из которых получается сама матрица прогноза X таким способом: $X_{ij} = r_i \cdot a_{ij} \cdot s_j$, где X_{ij} – элемент матрицы X и $a_{ij} = a_{ij}^{base}$ – элемент матрицы $A = A^{base}$, которая является матрицей «затрат-выпуска» за базовый год. Эти методы поэтому называются *бипропорциональными*. Все элементы этих векторов должны быть больше нуля, так как элементы матрицы A^{base} не должны менять знак при прогнозе.

- **Для метода *RAS***

Пусть вектор-столбец \vec{u} и вектор строка \vec{v} – вектора, содержащие, соответственно, окаймляющие суммы строк и столбцов матрицы за прогнозируемый год ($g + t$).

Теперь положим начальные значения всех элементов искомых векторов \vec{r} и \vec{s} равные единице.

Так как в итоге должно выполняться следующее равенство для всех сумм строк матрицы прогноза X :

$$(2) \quad u_i = \sum_j r_i x_{ij} s_j \Rightarrow r_i = \frac{u_i}{\sum_j x_{ij} s_j}$$

Также должно будет выполняться равенство для всех сумм столбцов матрицы прогноза X :

$$(2) \quad v_j = \sum_i r_i x_{ij} s_j \Rightarrow s_j = \frac{v_j}{\sum_i r_i x_{ij}}$$

Теперь мы знаем, как связаны вектора \vec{r} и \vec{s} , и можем запустить итеративный процесс их вычисления. Изначально, как было сказано выше, \vec{r} и \vec{s} – единичные вектора. Одна итерация устроена описанным ниже образом.

Сначала вычисляем все элементы вектора \vec{r} по формуле (2), где s_j – это либо j -й элемент вектора \vec{s} , который был вычислен во время предыдущей итерации алгоритма, либо, если это первая итерация, то $s_j = 1$, так как изначально вектор \vec{s} у нас единичный.

Затем вычисляем все элементы вектора \vec{s} по выведенной выше формуле (2), где r_i – это i -й элемент вектора \vec{r} , который был вычислен на предыдущем шаге этой же итерации. Теперь мы зацикливаем процесс.

Но, очевидно, нужен какой-нибудь критерий остановки алгоритма. Соответственно, после каждой итерации вычисляется промежуточная матрица A^{prom} по уже знакомой формуле: $A_{ij}^{prom} = r_i \cdot a_{ij} \cdot s_j$. Затем для нее вычисляются суммы строк и столбцов и берется

максимум из модуля отклонений получившихся сумм от тех, что записаны в векторах \vec{u} и \vec{v} . Когда это максимальное отклонение становится меньше некоторого $0 < \varepsilon \ll 1$, процесс останавливается, и по получившимся векторам \vec{r} и \vec{s} вычисляется финальная матрица X .

- Для метода **GRAS**

Этот метод отличается от своего «родителя» способностью построения матриц с отрицательными элементами. Целевая функция этого метода имеет следующий вид [Temurshoev, Webb, Yamano, 2011]:

$$(3) \quad \min_{z_{ij}} f(Z) = \sum_i \sum_j |a_{ij}| \cdot \left(z_{ij} \cdot \ln \left(\frac{z_{ij}}{e} \right) + 1 \right),$$

где $z_{ij} \equiv x_{ij}/a_{ij}$, если $a_{ij} \neq 0$, и $z_{ij} = 1$, иначе; $Z \equiv \{z_{ij}\}$; e – основание натурального логарифма.

Для начала представим матрицу $A = A^{base}$ как разность двух матриц таким образом: $A^{base} = P - N$, где P – матрица, содержащая все положительные элементы матрицы A^{base} и нули на местах неположительных элементов, а N – матрица, содержащая модули отрицательных элементов матрицы A^{base} и нули на местах остальных элементов. Определим, что для отрицательных элементов матрицы A^{base} соответствующие элементы матрицы X будут строиться таким образом: $x_{ij} = r_i^{-1} \cdot A_{ij}^{base} \cdot s_j^{-1}$.

Теперь запишем уравнение, отражающее равенство сумм строк нашего прогноза и элементов вектора \vec{u} , и преобразуем его:

$$(4) \quad \begin{aligned} u_i &= \sum_j r_i p_{ij} s_j - \sum_j r_i^{-1} n_{ij} s_j^{-1} \Rightarrow \\ u_i r_i &= r_i^2 \sum_j p_{ij} s_j - \sum_j n_{ij} s_j^{-1} \Rightarrow \\ r_i^2 \sum_j p_{ij} s_j - u_i r_i - \sum_j n_{ij} s_j^{-1} &= 0 \end{aligned}$$

Теперь решим (3) как квадратное уравнение относительно r_i :

$$r_i = \frac{u_i \pm \sqrt{u_i^2 + 4 \sum_j p_{ij} s_j \sum_j n_{ij} s_j^{-1}}}{2 \sum_j p_{ij} s_j},$$

где p_{ij} и n_{ij} – элементы матриц P и N соответственно.

Если мы оставим знак «минус» в числителе (4), то r_i будет меньше нуля, а так как все элементы вектора \vec{r} должны быть больше нуля, то нам подходит только решение со знаком «плюс».

К тому же коэффициент в квадратном уравнении при r_i^2 может быть равен нулю. Тогда получаем следующее уравнение и решение:

$$-u_i r_i - \sum_j n_{ij} s_j^{-1} = 0 \Rightarrow r_i = -\frac{\sum_j n_{ij} s_j^{-1}}{u_i}$$

Наконец, введем обозначения для удобства записи формул и запишем окончательную формулу для вычисления элементов вектора \vec{r} :

$$p_i(s) = \sum_j p_{ij} \cdot s_j, p_j(r) = \sum_i r_i \cdot p_{ij}, n_i(s) = \sum_j n_{ij} \cdot s_j^{-1} \text{ и } n_j(r) = \sum_i n_{ij} \cdot r_i^{-1}.$$

Теперь запишем общую формулу для вычисления r_i :

$$r_i = \begin{cases} \frac{u_i + \sqrt{u_i^2 + 4p_i(s)n_i(s)}}{2p_i(s)}, & p_i(s) > 0 \\ -\frac{n_i(s)}{u_i}, & p_i(s) = 0 \end{cases}$$

Аналогично получаем общую формулу для вычисления s_j :

$$s_j = \begin{cases} \frac{v_j + \sqrt{v_j^2 + 4p_j(r)n_j(r)}}{2p_j(r)}, & p_j(r) > 0 \\ -\frac{n_j(r)}{u_j}, & p_j(r) = 0 \end{cases}$$

Дальше строим итеративный алгоритм, аналогичный описанному для метода RAS, с тем же условием остановки.

В следующих двух методах целевая функция имеет квадратичный вид. Поэтому задача минимизации решалась стандартными средствами программного пакета Optimization Tools в MATLAB.

Метод *INSD*

Метод минимизации улучшенного нормализованного квадрата разности – Improved Normilzaed Square Difference, *INSD* – восходит к [Friedlander, 1961]. Целевая функция в этом случае принимает вид:

$$f = \sum_i \sum_j \frac{(x_{ij} - a_{ij})^2}{a_{ij}}$$

Через введенные выше переменные z_{ij} она выражается как

$$(5) \quad f(Z) = \sum_i \sum_j |a_{ij}| \cdot (z_{ij} - 1)^2$$

Как и раньше, предполагается, что при переходе от матрицы A к матрице X знаки элементов не могут измениться на противоположные, т.е. минимизация проводится при дополнительных ограничениях $z_{ij} \geq 0$. Кроме того, поскольку заданы окаймляющие итоги матрицы X , должны выполняться ограничения в виде равенств

$$u_i = \sum_j x_{ij} \quad \text{и} \quad v_j = \sum_i x_{ij}$$

II. Метод Куроды

Описанный ниже метод предложен в [Kuroda, 1988]. Целевая функция здесь принимает вид

$$f = \frac{1}{2} \cdot \sum_i \sum_j \left[\left(\frac{x_{ij}}{u_i} - r_{ij} \right)^2 \cdot w_{ij} + \left(\frac{x_{ij}}{v_j} - c_{ij} \right)^2 \cdot v_{ij} \right],$$

где $r_{ij} = a_{ij}/u_i^0$, $c_{ij} = a_{ij}/v_j^0$ для окаймляющих итогов u_i^0 и v_j^0 матрицы A , а «веса» w_{ij} и v_{ij} зависят от модификации метода, см. ниже. Записанная через переменные z_{ij} , эта функция принимает вид

$$f = \frac{1}{2} \cdot \sum_i \sum_j a_{ij}^2 \cdot \left[\left(\frac{z_{ij}}{u_i} - \frac{1}{u_i^0} \right)^2 \cdot w_{ij} + \left(\frac{z_{ij}}{v_j} - \frac{1}{v_j^0} \right)^2 \cdot v_{ij} \right].$$

Следуя [Temurshoev, Webb, Yamano, 2011], мы рассматриваем три модификации метода, различающиеся формулами для весов w_{ij} и v_{ij} :

Kuroda 1 – вариант «равнопроцентных весов» из [Kuroda, 1988, Case (2)]

$$w_{ij} = 1/r_{ij}^2, \quad v_{ij} = 1/c_{ij}^2;$$

Kuroda 2 – вариант с формулами весов

$$w_{ij} = \frac{u_i^2}{2}, \quad v_{ij} = \frac{v_j^2}{2}$$

из [Wilcohen, 1989], и

Kuroda 3 – вариант с тривиальными весами $w_{ij} = v_{ij} = 1$.

3. Источники данных для вычислительных экспериментов

Источниками данных для вычислительных экспериментов служат таблицы из базы данных WIOD (www.wiod.org). При это мы следили за тем, чтобы выполнялись следующие два условия: во-первых, таблица должна быть составлена на основе данных национальной статистической службы, а не построена как проекция таблицы за другой год. Здесь мы пользовались сведениями из [Egumban et al., 2012] об источниках данных WIOD. Во-вторых, мы включали таблицу в нашу базу данных для эксперимента только в том случае, когда были достаточные основания полагать, что эта таблица является базовой, т.е. составлена на основе первичной статистики, а не построена соответствующей национальной статистической службой как (уточненная) проекция базовой. Выбранные в результате годы по каждой стране помечены в следующей таблице знаком +.

Таблица 2. Наличие базовых таблиц для стран базы WIOD

Страны/Годы	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11
Австралия		+							+	+							
Австрия																	
Бельгия	+					+											
Бразилия	+	+				+				+							
Болгария	+																
Великобритания				+						+				+			
Венгрия						+					+					+	
Германия						+					+						+
Греция											+					+	
Дания			+														
Индия			+						+					+			
Индонезия	+					+					+						
Ирландия	+					+					+						
Испания	+					+					+						
Италия	+					+					+						
Канада			+														
Китай			+					+						+			
Кипр																	
Корея	+					+					+			+			
Латвия																	
Литва																	
Люксембург						+					+					+	
Мальта						+	+										
Мексика									+								
Нидерланды						+					+					+	
Польша						+					+						
Португалия							+					+					
Россия	+																
Румыния						+						+					
Словакия																	
Словения						+								+			
США			+					+					+				
Тайвань							+					+					
Турция				+				+									
Финляндия	+					+											
Франция																	
Чехия	+					+					+					+	
Швеция	+					+					+						
Эстония																	
Япония	+					+					+						+

Для большинства стран источниками информации о том, какие таблицы являются базовыми, являются сайты национальных статистических служб, где размещены и сами таблицы, и информация о них. Для некоторых стран соответствующей информации найти не удалось. Для *Кипра* нет таблиц за данный период. Для *Мексики* составлена лишь одна таблица за 2003 г., поэтому она не используется в данной работе, аналогичная ситуация наблюдается с *Канадой*, имеющей только одну базовую таблицу за 1997 г., и с *Россией* с единственной базовой таблицей за 1995 г. С *Нидерландами* и *Люксембургом* ситуация противоположная: насколько нам известно (см. источники сведений ниже), в этих странах базовые таблицы составляются ежегодно. Поскольку по другим странам мы рассматриваем в основном проекцию на срок около 5 лет, мы приняли решение ограничиться рассмотрением таблиц по Нидерландам и Люксембургом за 2000 и 2005 гг.

Не удалось найти точных сведений по *Австрии, Дании, Латвии, Литве, Словакии, Франции и Эстонии*.

Австралия

Сведения взяты с сайта Бюро статистики Австралии (Australian Bureau of Statistics <http://www.abs.gov.au/AusStats/ABS@.nsf/MF/5209.0.55.001>) в разделе Национальных счетов (National Accounts) информация о таблицах «затрат-выпуска» (Input-Output tables).

Бельгия

Сведения из [Avonds, 2005].

Бразилия

Сведения взяты с сайта Института географии и статистики Бразилии (IBGE – Instituto Brasileiro de Geografia e Estatística http://www.ibge.gov.br/english/estatistica/economia/matrizinsumo_produto/default.shtm) в разделе Economy ->System of National Accounts->Input-Product Matrix .

Болгария

На сайте Национального статистического института Болгарии (Hungarian Central Statistical Office <http://www.nsi.bg/en>) есть данные по таблицам ресурсов-использования (раздел Statistical data->Macroeconomic statistics->Supply Use Tables).

Великобритания

Информация взята с публикации на официальном сайте Национальной статистики Великобритании (UK Office for National Statistics <http://www.ons.gov.uk/ons/index.html>) [Wild, 2013], а также из [Mahajan, 2006].

Венгрия

Сведения взяты с сайта Центрального Статистического бюро Венгрии (https://www.ksh.hu/national_accounts_gdp), тема National Accounts .

Германия

Сведения по Германии любезно предоставила Ева Шварц (Eva Schwarz) из Федерального статистического бюро Германии.

Греция

Сведения из [Aroche-Reyes et al., 2012].

Индия

Сведения с сайта Indiatat (<http://www.indiastat.com/default.aspx>), раздел Экономика (Economy), Затраты – Выпуск (Input-Output), Input-Output Transactions.

Индонезия

Сведения из [Supriyanto et al., 2011].

Ирландия

Сведения из [APO Staidrimh, 2013].

Испания

Сведения взяты с сайта INE (Instituto Nacional de Estadística <http://www.ine.es/en/welcome.shtml>), раздел Экономика (Economy), Национальные счета (National Accounts).

Италия

Сведения из [Di Carlo, Santarelli, 2010].

Канада

Сведения взяты с сайта Статистики Канады (Statistics Canada <http://www.statcan.gc.ca/start-debut-eng.html>), раздел Economic Accounts, Input-Output Accounts.

Китай

Сведения из [Wu et al., 2015].

Кипр

Сведения из [Erumban et al., 2012].

Корея

Сведения из [Han'guk Ŭnhaeng, 2010].

Люксембург

Как нам любезно разъяснил Меран Кафай (Mehran Kafai), базовые (первичные) таблицы по Люксембургу составляются ежегодно.

Мальта

Сведения из [Erumban et al., 2012].

Мексика

Сведения из [Erumban et al., 2012].

Нидерланды

Сведения взяты из публикаций Statistics Netherlands по национальным счетам (National accounts of the Netherlands <http://www.cbs.nl/en-GB/menu/themas/macroeconomie/publicaties/publicaties/archief/2015/default.htm>) за 2002–2013 гг. Более подробную информацию о построении таблиц найти не удалось.

Польша

Сведения взяты с сайта Центрального статистического офиса Польши (Central Statistical Office of Poland <http://stat.gov.pl/en/>), раздел Национальные счета (National Accounts).

Португалия

Информация взята из публикации [Reis, Rua, 2009].

Россия

Сведения взяты из публикации [Erumban et al., 2012].

Румыния

Сведения из [Zaman et al., 2010].

Словения

Сведения взяты из двух источников: [Sila, Juvancic, 2005] и документа с сайта Статистического офиса Республики Словения (Statistical Office RS <http://www.stat.si/statweb>) «Methodological explanations».

Соединенные Штаты Америки

Сведения взяты с сайта Бюро экономического анализа (Bureau of Economical Analysis <http://www.bea.gov/index.htm>), раздел Industry, Benchmark Input-Output Accounts.

Тайвань

Сведения взяты с сайта Национальной статистики Китая (Тайвань) (National Statistics Republic of China (Taiwan)), раздел Статистика (Statistics from Statistical Bureau <http://eng.stat.gov.tw/mp.asp?mp=5>), I/O Tables, Statistical Tables.

Турция

Сведения взяты с сайта Турецкого института статистики (Turkish Statistical Institute), раздел Статистика по темам (Statistics by themes), Input-Output Tables.

Финляндия

Сведения взяты из публикации с сайта Статистики Финляндии (Statistics Finland) [OSF, 2003].

Чехия

Сведения из [Sixta, 2013].

Швеция

Сведения любезно сообщила нам Ёльва Петерсон Стрид (*Ylva Petersson Strid*) из Статистики Швеции.

Япония

Сведения взяты с сайта Министерства внутренних дел и коммуникаций Японии (Ministry of Internal Affairs and communications <http://www.soumu.go.jp/>), раздел Director-General for Policy Planning (Statistical Standards), Input-Output Tables for Japan.

4. Результаты расчетов

На основе полученных данных о базовых годах были проведены следующие расчеты: для каждой страны, имеющей больше одного базового года в рассматриваемом промежутке, для таблиц в базовых ценах и ценах покупателей были спрогнозированы первый, второй квадранты по отдельности и первый и второй квадранты вместе с базового года на следующий соседний базовый год для всех таких возможных пар. Например, у Германии базовые годы следующие: 2000, 2005, 2011, поэтому имеется две пары соседних базовых годов 2000->2005 и 2005->2011, так что было построено 12 прогнозов (по шесть на каждую пару: по три на пару для таблицы в базовых ценах и по три на пару для таблицы в ценах покупателей).

В итоге описанные прогнозы были построены для 52 пар базовых годов. В результате каждого из прогнозов получалась таблица следующего вида (примеры – для данных по

Голландии (I&II квадранты в ценах покупателей для прогнозов 1995->2000 и 2000->2005) и Испании (I, II и I&II квадранты в базовых ценах для прогноза 2000->2005):

Таблица 3. Нидерланды, объединенные I&II квадранты в ценах покупателей

1995 -> 2000

		Метрика														
Кв.	Метод	MAPE	R	SWAD	R	WAPE	R	PsiStat	R	RSQ	R	Inac	R	NO	R_all	CmR
I&II	GRAS	21,991	2	0,054	2	10,276	2	0,104	2	0,9967	2	3,8E+01	3	37	5	2
	Kuroda1	20,623	1	0,045	1	8,965	1	0,091	1	0,9976	1	8,9E-09	1	37	12	1
	INSD	22,503	3	0,058	3	10,807	3	0,109	3	0,9961	3	3,6E-08	2	37	1	3

Таблица 4. Нидерланды, объединенные I&II квадранты в ценах покупателей

2000 -> 2005

		Метрика														
Кв.	Метод	MAPE	R	SWAD	R	WAPE	R	PsiStat	R	RSQ	R	Inac	R	NO	R_all	CmR
I&II	GRAS	38,347	1	0,049	1	13,698	1	0,136	1	0,9944	1	6,1E+02	3	52	10	1
	Kuroda1	38,999	3	0,059	3	14,110	3	0,140	3	0,9931	3	6,3E-09	1	52	2	3
	INSD	38,613	2	0,050	2	13,871	2	0,138	2	0,9943	2	3,7E-08	2	52	6	2

Таблица 5. Испания, прогнозы в основных ценах

2000 -> 2005

		Метрика														
Кв.	Метод	MAPE	R	SWAD	R	WAPE	R	PsiStat	R	RSQ	R	Inac	R	NO	R_all	CmR
I	GRAS	29,347	2	0,081	2	11,475	2	0,113	2	0,9969	2	1,5E+02	3	42	5	2
	Kuroda1	28,583	1	0,024	1	10,145	1	0,100	1	0,9987	1	4,5E-09	1	42	12	1
	INSD	30,571	3	0,101	3	12,595	3	0,125	3	0,9951	3	2,3E-08	2	42	1	3
II	GRAS	33,325	1	0,010	1	3,757	1	0,039	1	0,9997	1	3,2E+02	1	6	12	1
	Kuroda1	64,781	2	0,367	2	30,592	2	0,305	2	0,9876	2	1,4E+05	2	0	6	2
	INSD	64,781	2	0,367	2	30,592	2	0,305	2	0,9876	2	1,4E+05	2	0	6	2
I&II	GRAS	30,385	3	0,056	2	9,686	2	0,097	2	0,9970	2	3,2E+02	3	48	4	2
	Kuroda1	29,222	1	0,026	1	8,281	1	0,083	1	0,9989	1	7,3E-09	1	48	12	1
	INSD	29,517	2	0,071	3	10,517	3	0,105	3	0,9953	3	4,8E-08	2	48	2	3

Строки таблицы соответствовали методам прогнозирования, а по столбцам стояли значения метрик на полученных прогнозах и ранг методов относительно этих значений.

Столбец R_all содержит разность между максимально возможной суммой рангов (она равна: $6 \cdot 3 = 18$) и наблюдаемой суммой рангов по всем метрикам, а столбец CmR – кумулятивный ранг на основе R_all .

В следующих таблицах приведено количество прогнозов, в которых каждый из методов был первым, вторым и третьим по рангу относительно каждой из метрик и по кумулятивному рангу. Эти таблицы приводятся отдельно по прогнозам в основных ценах и в ценах покупателей, а также отдельно по первому квадранту, второму квадранту и по объединенной матрице первого и второго квадрантов.

4.1. Таблицы в базовых ценах

а. Прогнозы первого квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	27	20	5
Second	20	4	28
Third	5	28	19

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	41	2	9
Second	11	2	39
Third	0	48	4

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	47	1	4
Second	5	3	44
Third	0	48	4

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	47	1	4
Second	5	3	44
Third	0	48	4

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	38	3	11
Second	14	4	34
Third	0	45	7

<i>Inac</i>	GRAS	Kuroda1	INSD
First	1	44	10
Second	5	2	42
Third	46	6	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	47	2	7
Second	5	3	41
Third	0	47	4

б. Прогнозы второго квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	19	21	18
Second	16	17	25
Third	17	14	9

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	39	10	3
Second	11	24	35
Third	2	18	14

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	39	9	4
Second	12	26	32
Third	1	17	16

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	39	9	4
Second	9	28	33
Third	4	15	15

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	35	12	6
Second	13	20	35
Third	4	20	11

<i>Inac</i>	GRAS	Kuroda1	INSD
First	18	27	7
Second	5	20	45
Third	29	5	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	34	15	5
Second	14	22	32
Third	4	15	15

в. Прогнозы первого и второго квадрантов вместе

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	18	23	11
Second	23	7	22
Third	11	22	19

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	32	7	13
Second	20	4	28
Third	0	41	11

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	44	3	5
Second	8	3	41
Third	0	46	6

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	38	4	10
Second	14	2	36
Third	0	46	6

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	39	6	7
Second	13	1	38
Third	0	45	7

<i>Inac</i>	GRAS	Kuroda1	INSD
First	0	47	6
Second	4	1	46
Third	48	4	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	40	5	13
Second	12	4	31
Third	0	43	8

4.2. Таблицы в ценах покупателя

а. Прогнозы первого квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	27	19	7
Second	21	4	26
Third	4	29	19

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	39	3	10
Second	13	3	37
Third	0	46	5

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	47	1	4
Second	5	4	44
Third	0	47	4

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	46	2	4
Second	6	4	43
Third	0	46	5

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	43	2	7
Second	9	5	39
Third	0	45	6

<i>Inac</i>	GRAS	Kuroda1	INSD
First	3	41	12
Second	4	5	40
Third	45	6	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	47	1	7
Second	5	7	39
Third	0	44	6

б. Прогнозы второго квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	18	23	18
Second	19	14	23
Third	15	15	11

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	35	13	4
Second	15	20	35
Third	2	19	13

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	41	8	3
Second	10	24	36
Third	1	20	13

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	40	10	2
Second	9	23	38
Third	3	19	12

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	37	10	6
Second	11	20	37
Third	4	22	9

<i>Inac</i>	GRAS	Kuroda1	INSD
First	18	28	7
Second	4	20	45
Third	30	4	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	38	12	5
Second	11	23	34
Third	3	17	13

в. Прогнозы первого и второго квадрантов вместе

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	15	24	13
Second	29	6	17
Third	8	22	22

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	32	8	12
Second	19	2	31
Third	1	42	9

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	45	3	4
Second	7	4	41
Third	0	45	7

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	40	3	9
Second	12	5	35
Third	0	44	8

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	41	6	5
Second	11	1	40
Third	0	45	7

<i>Inac</i>	GRAS	Kuroda1	INSD
First	0	48	7
Second	4	0	45
Third	48	4	0

<i>CmR</i>	GRAS	Kuroda1	INSD
First	39	6	11
Second	13	3	33
Third	0	43	8

4.3. Таблицы, которые не могут быть сбалансированы

Некоторые таблицы не могут быть сбалансированы точно методом RAS или GRAS (например, в случае разреженных матриц, см. [Miller, Blair, 2009, 7.4.9]). Иногда таблица вообще не может быть сбалансирована точно при соблюдении условия постоянства знаков, поскольку, например, иногда в СНС начинают учитывать новые категории, которым в таблице за предыдущий (базовый) год отвечают нулевые элементы. Препятствием для точной балансировки может также послужить изменение знака в какой-либо категории конечного спроса на противоположный. Поэтому может представлять интерес построение приближенной таблицы, в которой наблюдаются (относительно небольшие) невязки.

Ниже перечисляются все случаи, в которых матрица прогноза первого квадранта не получилась достаточно сбалансированной, т.е. когда значение *Inac* (максимум абсолютных величин невязок между суммами строк и столбцов получившейся матрицы и заданными окаймляющими итогами) не менее 10). Они даны отдельно для таблиц в базовых ценах и в ценах покупателя. Из-за наличия отрицательных элементов во втором квадранте и из-за того, что матрица второго квадранта обычно разрежена, при балансировке отдельно второго и одновременно первого и второго квадрантов такие случаи более распространены, и полный список мы здесь не приводим. В целом, оценки сравнительной эффективности методов, отраженные в приведенных выше таблицах, не меняются, если ограничить статистическую базу только достаточно сбалансированными таблицами (с *Inac* < 10).

а. Прогнозы с $I_{нас} > 10$ (для квадранта I таблиц в основных ценах)

Страна	Годы	Методы		
BEL	1995 -> 2000	GRAS		
HUN	2000 -> 2005	GRAS		
IND	1998 -> 2003	GRAS		
IDN	1995 -> 2000	GRAS	Kuroda	INSD
IDN	2000 -> 2005	GRAS		
IRL	2000 -> 2005	GRAS		
ESP	1995 -> 2000	GRAS		
ESP	2000 -> 2005	GRAS		
CHN	1997 -> 2002	GRAS	Kuroda	INSD
CHN	2002 -> 2007	GRAS	Kuroda	INSD
KOR	1995 -> 2000	GRAS		
KOR	2000 -> 2005	GRAS		
MLT	2000 -> 2001	GRAS		
NLD	2000 -> 2005	GRAS		
NLD	2005 -> 2010	GRAS		
POL	2000 -> 2005	GRAS		
ROU	2000 -> 2006	GRAS	Kuroda	INSD
USA	1997 -> 2002	GRAS		
TWN	2001 -> 2006	GRAS		
TUR	1998 -> 2002	GRAS	Kuroda	INSD
FIN	1995 -> 2000	GRAS		
CZE	1995 -> 2000	GRAS		
CZE	2000 -> 2005	GRAS		
CZE	2005 -> 2010	GRAS		
SWE	1995 -> 2000	GRAS		
SWE	2000 -> 2005	GRAS		
JPN	1995 -> 2000	GRAS		
AUS	1996 -> 2003	GRAS	Kuroda	INSD
BRA	2000 -> 2005	GRAS		
BGR	2000 -> 2005	GRAS		
GBR	1998 -> 2004	GRAS		

б. Прогнозы с $I_{нас} > 10$ (для квадранта I таблиц в ценах покупателей)

Страна	Годы	Методы		
BEL	1995 -> 2000	GRAS		
HUN	2000 -> 2005	GRAS		
IND	1998 -> 2003	GRAS		
IDN	1995 -> 2000	GRAS	Kuroda	INSD
IDN	2000 -> 2005	GRAS		
IRL	2000 -> 2005	GRAS		
ESP	1995 -> 2000	GRAS	Kuroda	INSD
ESP	2000 -> 2005	GRAS		
CHN	1997 -> 2002	GRAS	Kuroda	INSD
CHN	2002 -> 2007	GRAS	Kuroda	INSD
KOR	1995 -> 2000	GRAS		
KOR	2000 -> 2005	GRAS		
MLT	2000 -> 2001	GRAS		
NLD	2000 -> 2005	GRAS		
NLD	2005 -> 2010	GRAS		
POL	2000 -> 2005	GRAS		
ROU	2000 -> 2006	GRAS	Kuroda	INSD
USA	1997 -> 2002	GRAS		
TWN	2001 -> 2006	GRAS		
TUR	1998 -> 2002	GRAS	Kuroda	INSD
FIN	1995 -> 2000	GRAS		
CZE	1995 -> 2000	GRAS		
CZE	2000 -> 2005	GRAS		
CZE	2005 -> 2010	GRAS		
SWE	1995 -> 2000	GRAS		
SWE	2000 -> 2005	GRAS		
JPN	1995 -> 2000	GRAS		
AUS	1996 -> 2003	GRAS	Kuroda	INSD
BRA	2000 -> 2005	GRAS		
BGR	2000 -> 2005	GRAS		
GBR	1998 -> 2004	GRAS		
IND	2003 -> 2008	GRAS		
ITA	2000 -> 2005	GRAS		

4.4. Сбалансированные результаты расчетов: исходные данные

Напомним, что мы условно считаем сбалансированными прогнозные таблицы, в которых величина $I_{нас}$ (максимальное абсолютное отклонение итога по столбцу или строке от экзогенно заданного) не больше 10. В зависимости от выбранного варианта таблицы (квадрант I, квадрант II или объединенные квадранты I и II) и от того, посчитаны они в базовых ценах или в ценах покупателей, получаем следующие списки сбалансированных прогнозов по разным странам и годам.

<i>Страна</i>	<i>Годы</i>	Базовые цены			Цены покупателя		
		I кв.	II кв.	I&II кв.	I кв.	II кв.	I&II кв.
AUS	2003 -> 2004	"+"	"+"	"+"	"+"	"+"	"+"
BRA	1995 -> 1996	"+"	"+"	"+"	"+"	"+"	"+"
BRA	1996 -> 2000	"+"		"+"	"+"		"+"
BGR	1995 -> 2000	"+"	"+"	"+"	"+"	"+"	"+"
GBR	2004 -> 2008	"+"			"+"		
HUN	2005 -> 2010	"+"	"+"	"+"	"+"		
DEU	2000 -> 2005	"+"			"+"		
DEU	2005 -> 2011	"+"		"+"	"+"		"+"
GRC	2005 -> 2011	"+"			"+"		
IND	2003 -> 2008	"+"					
IRL	1995 -> 2000	"+"	"+"	"+"	"+"	"+"	"+"
ITA	1995 -> 2000	"+"			"+"		
ITA	2000 -> 2005	"+"		"+"			
KOR	2005 -> 2008	"+"	"+"	"+"	"+"	"+"	"+"
PRT	2001 -> 2006	"+"	"+"	"+"	"+"	"+"	"+"
SVN	2000 -> 2008	"+"			"+"		
USA	2002 -> 2007	"+"			"+"		
JPN	2000 -> 2005	"+"	"+"	"+"	"+"	"+"	"+"
JPN	2005 -> 2011	"+"		"+"	"+"		"+"
LUX	2000 -> 2005	"+"			"+"		
LUX	2005 -> 2010	"+"		"+"	"+"		"+"
BRA	2000 -> 2005		"+"			"+"	

Таким образом, остается всего одна страна, для которой прогнозы для всех пар базовых лет получились достаточно сбалансированными. Это Португалия.

**4.5. Результаты расчетов для сбалансированных таблиц:
таблицы в базовых ценах**

а. Прогнозы первого квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	16	4	1
Second	4	0	17
Third	1	17	3

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	18	1	2
Second	3	0	18
Third	0	20	1

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	18	0	3
Second	3	1	17
Third	0	20	1

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	18	0	3
Second	3	1	17
Third	0	20	1

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	14	1	6
Second	7	0	14
Third	0	20	1

б. Прогнозы второго квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	2	4	3
Second	4	1	4
Third	3	4	2

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	4	4	1
Second	5	0	4
Third	0	5	4

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	6	1	2
Second	3	3	3
Third	0	5	4

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	5	3	1
Second	2	3	4
Third	2	3	4

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	3	3	3
Second	5	1	3
Third	1	5	3

в. Прогнозы первого и второго квадрантов вместе

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	7	4	2
Second	5	1	7
Third	1	8	4

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	9	2	2
Second	4	3	6
Third	0	8	5

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	12	0	1
Second	1	1	11
Third	0	12	1

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	9	0	4
Second	4	1	8
Third	0	12	1

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	9	2	2
Second	4	0	9
Third	0	11	2

4.6. Результаты расчетов для сбалансированных таблиц: таблицы в базовых ценах

а. Прогнозы первого квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	15	4	0
Second	3	0	16
Third	1	15	3

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	15	2	2
Second	4	0	15
Third	0	17	2

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	17	0	2
Second	2	1	16
Third	0	18	1

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	17	0	2
Second	2	1	16
Third	0	18	1

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	16	0	3
Second	3	0	16
Third	0	19	0

б. Прогнозы второго квадранта

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	2	3	3
Second	3	1	4
Third	3	4	1

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	3	4	1
Second	4	0	4
Third	1	4	3

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	3	3	2
Second	5	0	3
Third	0	5	3

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	2	5	1
Second	4	0	4
Third	2	3	3

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	2	3	3
Second	5	1	2
Third	1	4	3

в. Прогнозы первого и второго квадрантов вместе

<i>MAPE</i>	GRAS	Kuroda1	INSD
First	5	3	3
Second	6	1	4
Third	0	7	4

<i>SWAD</i>	GRAS	Kuroda1	INSD
First	8	2	1
Second	3	1	7
Third	0	8	3

<i>WAPE</i>	GRAS	Kuroda1	INSD
First	10	0	1
Second	1	1	9
Third	0	10	1

<i>PsiStat</i>	GRAS	Kuroda1	INSD
First	8	0	3
Second	3	1	7
Third	0	10	1

<i>RSQ</i>	GRAS	Kuroda1	INSD
First	7	2	2
Second	4	0	7
Third	0	9	2

Краткий анализ отраженных в этих таблицах результатов дан в Заключение.

5. Заключение

Для сравнительного анализа методов проекции мы используем результаты лишь сбалансированные таблицы из результатов вычислительных экспериментов, описанные в пп.4.4, 4.5 и 4.6. По нашим данным, из трех рассматриваемых методов проекции наиболее эффективным показал себя пропорциональный метод GRAS. Преимущество этого метода безусловно по результатам проекции первого квадранта таблицы использования (матрица промежуточного использования), на котором метод сводится к методу RAS, а также по результатам проекции первого и второго квадранта вместе. Этот эффект наблюдается при проекции таблиц как в ценах покупателей, так и в основных ценах. Тем не менее при проекции таблицы конечного спроса (квадрант II) преимущество этого метода выглядит

бесспорным, но не подавляющим: для таблиц в ценах покупателей по основным метрикам он примерно в четырех случаях из пяти оказывается более эффективным, чем следующий за ним метод Куроды. Для таблиц в основных ценах недостаток статистики не позволяет сделать определенных выводов, но предварительно можно утверждать, что здесь преимущество GRAS над квадратичными методами несколько менее выражено.

Отметим, что наши результаты повторяют на таблицах из проекта WIOD упомянутые во Введении фрагменты расчетов по Испании и Нидерландам из [Temurshoev, Webb, Yamano, 2011] (см. табл. 4 и 5). Однако наше исследование показывает, что обнаруженные там случаи относительной эффективности квадратичных методов по сравнению с GRAS достаточно редки и могут считаться статистической флуктуацией.

Полученное очередное подтверждение эффективности пропорциональных методов при проекции таблиц «затраты – выпуск» служит аргументом в пользу того, чтобы при построении таблиц использования по России за 2000-е годы использовать именно такие методы.

При проекции таблицы промежуточного использования (квадрант I), а также всей таблицы использования (объединенных квадрантов I и II) как в основных ценах, так и в ценах покупателей на втором месте по сравнительной эффективности оказался метод INSD. Возможно, причина достаточно высокой сравнительной эффективности этого метода чисто математическая. Как отмечено в [Huang et al., 2008], целевая функция (5) этого метода представляет собой аппроксимацию Тейлора второго порядка аналогичной функции (3) метода GRAS при $z_{ij} = 1$, поэтому естественно, что построенные этими методами матрицы близки (особенно при проекции на небольшие временные промежутки в 2–5 лет, когда изменения коэффициентов a_{ij} малы, т.е. значения z_{ij} близки к 1).

Относительно низкая эффективность метода Куроды оказалась для нас сюрпризом. Вероятно, это означает, что наблюдаемая раньше относительно высокая эффективность этого метода на таблицах использования основана на случайных флуктуациях. Возможно, здесь играет роль какая-то особенность агрегации национальных таблиц в таблицы WIOD. Мы видим, тем не менее, что этот метод относительно эффективен при измерении результатов метрикой MAPE, что позволяет рекомендовать построенные на его основе прогнозы в задачах, в которых важны лишь относительные, а не абсолютные значения изменений отдельных показателей.

Наша методика сравнительного исследования может быть улучшена при дальнейших исследованиях для получения более точных результатов. Одна из возможных причин неточности в наших расчетах – заметное количество несбалансированных таблиц. Более тонкая работа с данными должна исключить большинство таких случаев и уточнить

результаты расчетов. Исключение импорта и экспорта (известных из официальных статистических данных) из второго квадранта может несколько изменить значения метрик на прогнозах по квадранту II. Еще одна причина возможной неточности (хотя и не главная) – зависимость результата применения квадратичных методов от выбранного алгоритма условной квадратичной оптимизации: мы обнаружили, что результаты несколько различаются при применении стандартных средств такой оптимизации, реализованных в системе MATLAB, и при применении более эффективного на этой задаче барьерного метода, реализованного в библиотеке GUROBI.

Вместе с тем обнаружилась еще одна область возможного применения квадратичных методов проекции Kuroda 1 и INSD. В ряде случаев таблица не может быть сбалансирована пропорциональным методом из-за значительного изменения ее структуры (особенно в случае разреженных матриц): в нашем вычислительном эксперименте, даже по первому квадранту (в котором метод RAS наиболее эффективен) это происходило в 31 случаях из 52 для расчетов в основных ценах и в 33 случаях из 52 – в ценах покупателей (см. п. 4.3). В то же время в 25 случаях из 31 (соответственно, в 27 из 33) методы Kuroda 1 и INSD давали достаточно удовлетворительный результат. Таким образом, в случае, когда метод RAS не дает сбалансированную матрицу, можно рекомендовать для проекции таблиц использования квадратичные методы: INSD и метод Куроды.

Литература

Баранов Э.Ф., Ким И.А., Пионтковский Д.И., Старицына Е.А. Вопросы построения таблиц «затраты – выпуск» России в международных классификаторах // Экономический журнал ВШЭ. 2014. Т. 18. № 1. С. 7–42.

Central Statistics Office, Supply and Use Tables for Ireland 2009–2013 <http://www.cso.ie/en/media/csoie/releasespublications/documents/economy/2009/supplyuse_2009.pdf>.

Aroche-Reyes F., García Muñiz A.S. Modelling economic structures from a Qualitative Input-Output Perspective: Greece in 2005 and 2010. 012.

Avonds L. Belgian Input-Output Tables: State of the Art, paper presented at the 15th International Conference on Input-Output Techniques, Beijing <<http://iioa.org/conferences/15th/pdf/avonds.pdf>>.

Di Carlo C., Santarelli E. Ministry of Economic Development, Department of Communications, Contribution of ICT to economic growth in Italy: Input Output analysis. November 2010 <http://www.sviluppoeconomico.gov.it/images/stories/pubblicazioni/Contribution_of_ICT_to_economic_.pdf>.

Erumban A.E., Gouma R., Timmer M., de Vries G., de Vries K. Sources for National Supply and Use Table Input files. World Input-Output Database (WIOD), 2012b <http://www.wiod.org/publications/source_docs/sut_input_sources.pdf>.

Eurostat. Eurostat Manual of Supply, Use and Input-Output Tables. 2008.

Friedlander D. A technique for estimating a contingency table, given the marginal totals and some supplementary data // *Journal of the Royal Statistical Society. Series A (General)*. 961. P. 412–420.

Gunluk-Senesen G., Bates J.M. Some Experiments With Methods of Adjusting Unbalanced Data Matrices // *Journal of the Royal Statistical Society*. 1988. Series A. 151. P. 473–490.

Han'guk Ũnhaeng. The Bank of Korea: A Sixty-year History. Bank of Korea, 2010. P. 135–136.

Huang W., Kobayashi S., Tanji H. Updating an input-output matrix with sign-preservation: some improved objective functions and their solutions // *Economic Systems Research*. 2008. Vol. 20 (1). P. 111–123.

Mahajan S. UK Input-Output Analysis // Office for National Statistics. L., 2006 <http://www.statistics.gov.uk/downloads/theme_economy/Input_Output_Analyses_2006_edition.pdf>.

Miller R.E., Blair P.D. Input-Output Analysis: Foundations and Extensions. Cambridge University Press, 2009.

Official Statistics of Finland (OSF) Methodological reform of national accounts: integration of supply and use tables – 2003 (2/2003).

Reis H., Rua A. An Input-Output Analysis: Linkages versus Leakages // *International Economic Journal*. 2009. Vol. 23. No. 4. P. 527–544.

Supriyanto, Buyung Airlangga, and Etjih Tasriah. Compiling the Supply and Use Tables: The Indonesia Experience // Conference on the Development of Supply and Use Tables in Asia and the Pacific, 12–13 July 2011, ADB Headquarters, Manila, Philippines <<http://adb.org/sites/default/files/INO-11-July-2011.pdf>>.

Sila U., Juvancic L. Regional input-output table: The case of Eastern Slovenia // *Jahrbuch der Österreichischen Gesellschaft für Agrarökonomie*. 2005. Vol. 13. P. 121–134.

Sixta J. Development of Input-Output Tables in the Czech Republic // *Statistika*. 2013. Vol. 93. No. 2. P. 4–14.

Stone R.A. Input-output accounts and national accounts // Organization for European Economic Cooperation. Paris, 1961.

Temurshoev U., Miller R.E., Bouwmeester M.C. A Note on the GRAS Method // *Economic Systems Research*. 2013. Vol. 25:3. P. 361–367.

Temurshoev U., Webb C., Yamano N. Projection of Supply and Use Tables: Methods and their Empirical Assessment // *Economic Systems Research*. 2011. Vol. 23. No. 1. P. 91–123.

Temurshoev U., Timmer M.P. Joint estimation of supply and use tables // *Papers in Regional Science*. 2011. Vol. 90. No. 4. P. 863–882.

Wild R. Commentary on Supply and Use balanced estimates of annual GDP, 1997–2011 // ONS – 31 July 2013 <<http://www.ons.gov.uk/ons/rel/input-output/input-output-supply-and-use-tables/commentary-on-supply-and-use-balanced-estimates-of-annual-gdp--1997-2013/index.html>>.

Wu H., Ito K. Reconstructing China's Supply-Use and Input-Output Tables in Time Series / Research Institute of Economy, Trade and Industry (RIETI), 2015.

Zaman, G., Surugiu, M., Surugiu C. Time-Stability of The Coefficients: An Input-Output Analysis On Romania Case. *Analele Stiintifice ale Universitatii Alexandru Ioan Cuza din Iasi-Stiinte Economice*. 2010. No. 57. P. 483–503.

Сведения об авторах:

Пионтковский Дмитрий Игоревич – д.ф.м.н., профессор департамента математики на факультете экономических наук НИУ ВШЭ; E-mail: dpiontkovski@hse.ru

Соколов Денис Дмитриевич – студент 4-го курса бакалавриата ОПМИ ФКН НИУ ВШЭ; E-mail: sokol0077@gmail.com

Старчикова Ольга Сергеевна – студентка 4-го курса бакалавриата ОПМИ ФКН НИУ ВШЭ; E-mail: starchikovaos@gmail.com

Piontkovski, D. I., Sokolov, D. D., Starchikova, O. S.

A Comparison of the Mathematical Projection Methods of IO accounts on the database of the WIOD project [Electronic resource] : Working paper WP2/2015/07/ ; National Research University Higher School of Economics. – Electronic text data (1 MB). – Moscow: Higher School of Economics Publ. House, 2015. – (Series WP2 “Quantitative Analysis of Russian Economy”) (in Russian). – 37 p.

We investigate the relative effectiveness of the projection methods of Supply and Use tables in relation to Use tables. The empirical bases of the study are the Use tables of 28 countries for the period from 1995 to 2010 from WIOD project. We conduct a comparative study of three mathematical methods that have proven the most effective in constructing projection of Use tables for Spain and the Netherlands from the investigation of [Temurshoev, Webb and Yamono, 2011]. The most effective of these methods is the biproportional GRAS method. The results of applying this method under the number of criteria are closer to the published tables than the results of the INSD method and Kuroda method, which are based on quadratic programming. We conclude that GRAS method is a priority in the extrapolation of Use tables for Russia.

Keywords: WIOD project, RAS, Kuroda method, supply and use tables

Classification JEL: D57, C53, C67

Препринт WP2/2015/07
Серия WP2
Количественный анализ в экономике

Пионтковский Дмитрий Игоревич, Соколов Денис Дмитриевич,
Старчикова Ольга Сергеевна

**Сравнение математических методов прогнозирования
таблиц «затраты – выпуск» на основе базы данных WIOD**

Зав. редакцией оперативного выпуска *А.В. Заиченко*
Технический редактор *Ю.Н. Петрина*

Изд. № 1964

Национальный исследовательский университет
«Высшая школа экономики»