

**ОБЪЕМ ИНФОРМАЦИИ В МАТРИЦАХ ПРЯМЫХ ЗАТРАТ
РАЗНОЙ РАЗМЕРНОСТИ В МОДЕЛИ ЛЕОНТЬЕВА
«ЗАТРАТЫ – ВЫПУСК»**

В. А. Пархименко¹⁾, А. А. Быков²⁾

¹⁾ *Белорусский государственный университет информатики и радиоэлектроники,
Минск, Беларусь, parkhimenko@bsuir.by*

²⁾ *Белорусский государственный экономический университет,
Минск, Беларусь, bikov_a@bseu.by*

Рассматривается проблема размерности матрицы прямых затрат в модели Леонтьева с точки зрения содержащегося в матрице объема информации о реальных экономических явлениях и процессах. Предлагается концептуальный подход к получению оценок объема информации в матрицах прямых затрат разного уровня детализации / агрегирования относительно «истинного» объема информации в экономической деятельности субъектов хозяйствования.

Ключевые слова: информация; межотраслевой баланс; модель Леонтьева; таблицы «затраты – выпуск»; экономическое моделирование.

**THE INFORMATION CONTENT OF DIRECT REQUIREMENTS
MATRICES OF VARYING DIMENSIONS IN THE LEONTIEF
INPUT-OUTPUT MODELS**

U. A. Parkhimenka^{a)}, A. A. Bykau^{b)}

^{a)} *Belarusian State University of Informatics and Radioelectronics,
Minsk, Belarus, parkhimenko@bsuir.by*

^{b)} *Belarusian State Economic University,
Minsk, Belarus, bikov_a@bseu.by*

This paper addresses the issue of the dimensionality of the direct requirements matrix in the Leontief model from the perspective of the amount of information it contains about real economic phenomena and processes. An approach is proposed for estimating the information content in direct requirements matrices of different levels of detail/aggregation relative to the ‘true’ amount of information in the economic activities of companies.

Keywords: information; input-output table; Leontief model; input-output analysis; economic modeling.

1. Введение

Матрица прямых в модели Леонтьева «затраты – выпуск» [1] представляет собой квадратную матрицу, в ij -й ячейке которой указывается величина прямых затрат продукции (товаров и услуг) i -й отрасли, потребляемой в j -й отрасли на единицу выпуска (1 рубль) продукции j -й отрасли.

Обычная практика для прикладных экономических расчетов использовать подобного вида матрицу размерностью от нескольких десятков до нескольких сотен отраслей. Например, в Республике Беларусь в настоящее время размерность таблиц «затраты – выпуск» (на основе которых рассчитывается матрица прямых затрат и другие компоненты модели Леонтьева) составляет 83×83 , в Российской Федерации – 99×99 , в США – 398×398 .

Очевидно, что более детализированные таблицы содержат значительно больше информации об экономических явлениях и процессах, наблюдаемых в национальной экономике за годовой период. Вместе с тем создание детализированных таблиц требует значительных расходов ресурсов и времени.

В этом контексте научный интерес вызывают оценки объема информации в матрицах прямых затрат разной размерности в модели Леонтьева «затраты – выпуск».

2. Оценка «истинного» объема информации

Предложим концептуальный подход к осмыслению и поиску решения для этой задачи.

Попробуем представить «истинный» объем информации, который «содержится» непосредственно в самих реально протекающих экономических процессах. Для этого от уровня отрасли (совокупности предприятий) перейдем к наименьшему ее компоненту – одному предприятию.

Пусть в данный момент в стране есть N организаций (компаний, предприятий, учреждений), которые, помимо выпуска конечной продукции и услуг для потребления, выпускают так называемые промежуточные товары и оказывают услуги, которые затем потребляются как ресурсы в других производственных процессах.

Если представить, что каждая организация осуществляет поставки своей продукции каждой другой организации, то этот процесс можно представить таблицей транзакций (потоков товаров и услуг) в стоимостном выражении размерностью $N \times N$. Число потенциальных транзакций, таким образом, будет равно N^2 .

Из этой таблицы путем деления (по каждому j -му столбцу) абсолютных значений на соответствующие выпуски продукции j -й отрасли

($j = \overline{1, N}$) можно перейти к матрице прямых затрат максимально возможной размерности (степени детализации), назовем ее условно «истинной».

Пусть в каждой ячейке матрицы мы можем записать число от 0 до 1 с точностью в шесть знаков после запятой. Тогда объем информации, который потенциально будет содержаться в одной ячейке, $-\log_2 1000001 \approx \approx 19,9 \approx 20$ бит. А во всей матрице – $20 \cdot N^2$ бит.

По данным Белстата, в 2024 году в Беларуси были зарегистрировано 130 463 коммерческих организаций, в том числе 11 384 малых организаций, 2 093 средних и 1 598 крупных, т.е. без учета так называемых микро-организаций – 15 075.

Отталкиваясь от общего числа коммерческих организаций и нереалистично допуская наличие транзакций между всеми компаниями, получим оценку информации, содержащейся в «истинной» матрице: $20 \cdot N^2 \approx \approx 20$ бит $\cdot 17,02$ млрд ячеек $\approx 39,49$ Гбайт.

При переходе к таблицам меньшей размерности, например, 83×83 , как это принято для составления таблиц «затраты – выпуск» в Республике Беларусь по числу видов экономической деятельности (ВЭД), получаем объем информации, равный $20 \cdot 83^2 = 20$ бит $\cdot 6889$ ячеек $\approx 16,82$ Кбайт, т.е. всего лишь $4 \cdot 10^{-5}\%$ от информации, содержащейся в истинной матрице.

Очевидно, однако, что в реальной действительности наша «истинная» матрица будет содержать огромное количество нулевых ячеек, так как ни одна из наугад выбранных организаций не будет взаимодействовать с каждой другой организацией без исключения, а лишь с некоторым их числом. Прежде всего потому, что технологический процесс отдельного производства требует только определенного набора входных ресурсов (например, для производства автотракторной техники нужен металл и пластмассы, но не продукции животноводства). Кроме того, в одно и то же время на рынке находится множество аналогичных компаний-конкурентов, которые, как правило, не работают с одними и теми же поставщиками или клиентами.

Вряд ли в настоящее время возможны точные оценки на этом «истинном» детализированном уровне матрицы прямых затрат, так как такие данные не собираются и не анализируются статистическими органами. (Возможно, при построении цифровой экономики и полному отражению всех экономических потоков и транзакций в централизованной информационной системе такие данные будут доступны государству как архитектору и регулятору цифровой экономики.) Неизвестны и статистические характеристики и распределения этих данных на этом уровне.

Предложим, однако, некоторый эвристический подход, который, с нашей точки зрения, позволят сделать хотя бы грубые и приблизительные оценки.

Будем исходить из того допущения, что можно охарактеризовать всех экономических субъектов (в данном случае коммерческие организации) средним числом поставщиков и клиентов – μ . Отталкиваясь от данных Белстата за 2024 год и предполагаемого числа поставщиков / клиентов у разного типа организаций, рассчитаем средневзвешенное значение параметра μ (табл. 1).

Таблица 1

Оценка параметра μ

Тип организаций	Количество субъектов	Предполагаемое среднее число поставщиков / клиентов (μ)
Микроорганизации (до 15 человек)	115 388	10
Малые организации (от 16 до 100 человек)	11 384	100
Средние организации (до 250 человек)	2 093	1 000
Крупные организации (свыше 250 человек)	1 598	10 000
Коммерческие организации, ВСЕГО	130 463	≈ 156

Если отталкиваться от числа 156 из табл. 1, означающего, что в среднем на одну организацию приходится 156 поставщиков/клиентов, то количество отличных от нуля ячеек в «истинной» матрице прямых затрат будет равно не N^2 , а $\mu \cdot N$ (в каждом столбце – μ , а во всей матрице – $\mu \cdot N$) и количество информации, соответственно, – $20 \cdot \mu \cdot N$ бит, что для случая белорусской экономики дает нам оценку в 49 688,1 Кбайт.

3. Оценка объема информации в матрицах прямых затрат разной размерности

Далее предположим, что матрицы, чья размерность во много раз меньше размерности «истинной» матрицы (в случае белорусской экономики мы полагаем, что речь может идти о размерности в несколько сотен отраслей) и примерно соответствует сложившейся статистической практике разработки таблиц «затраты – выпуск», – для таких матриц ввиду их существенной агрегированности объем информации может быть оценен уже по первоначальной формуле, т.е. как $\approx 20 \cdot N^2$, так как количество нулевых ячеек будет крайне незначительным.

В табл. 2 сведены расчетные оценки объема информации в матрицах прямых затрат в модели Леонтьева при разной их размерности и при указанном выше допущении.

Таблица 2

Оценка объема информации в матрицах разной размерности

Размерность матрицы (N)	Потенциальный объем информации, Кбайт	Объем информации от «истинного» объема
130 463	49 688,1	100,0%
500	610,4	1,23%
250	152,6	0,31%
100	24,4	0,05%
83	16,8	0,03%

Полученное значение 0,03% информации, содержащейся в отчетственных таблицах «затраты – выпуск» (точнее говоря, в таком элементе, как матрица прямых затрат), на первый взгляд, вряд ли может рассматриваться как достаточное. И даже матрицы размерностью около 500 строк/столбцов не позволяют говорить о «репрезентации» большого объема информации по отношению к «истинной» матрице прямых затрат.

С другой стороны, экономическая наука часто оперирует такими агрегированными показателями, как ВВП, индекс потребительских цен, уровень безработицы, которые представляют собой просто скалярные значения. Использование системы взаимосвязанных количественных показателей, например, в форме модели Леонтьева «затраты – выпуск», в любом случае представляет собой шаг в сторону более полного описания реальных экономических явлений и процессов.

4. Заключение

Полученные оценки (см. табл. 2), конечно, являются очень грубыми и предварительными.

Остаются не более чем экспертными предположениями такие «логические звенья» рассмотренного выше подхода, как оценка параметра $\mu = 156$, так и утверждение, что для матриц размерностью меньше 1000 количество нулевых ячеек пренебрежительно мало.

Видны и направления для развития и уточнения предложенного подхода.

Агрегирование таблиц меньшей размерности из таблиц больше размерности получается путем суммирования соответствующих ячеек. Это может приводить к необходимости отведения большего количества бит на хранение информации в каждой ячейке. Таким образом, эта тенденция

будет противодействовать (как представляется, незначительно) общему уменьшению объема информации в таблицах меньшей размерности.

Следует учесть информацию об импортных товарах и услугах.

Кроме того, таблицы «затраты – выпуск» содержат и другие компоненты: конечный спрос по категориям институциональных единиц (домохозяйства, государственные учреждения и т.д.).

Крайне важный вопрос (не затрагиваемый в данном исследовании) критерия построения самих таблиц: по отраслевому принципу или продуктовому.

Наконец, важно учесть статистическое распределение значений в матрице прямых затрат и оценить информационную энтропию по Шеннону [2]. Даже если в качестве эвристического предположения можно предположить, что те 156 значений в каждом столбце (или строке) матрицы прямых затрат имеют распределение Парето (это довольно традиционное предположение для многих процессов в экономике), то вполне можно утверждать, что объем информации, измеряемый как информационная энтропия по Шеннону, будет для «истинного» случая значительно меньше, чем указано в табл. 2.

Библиографические ссылки

1. *Miller R. E., Blair P. D.* Input-output analysis: foundations and extensions. Cambridge university press, 2022.
2. *Shannon C. E.* A Mathematical Theory of Communication // Bell System Technical Journal. 1948. Vol. 27, iss. 3. P. 379–423.