

Модели комплексирования квартальных данных на основе структурных характеристик эталонной матрицы

Владимир Ильич Моторин^{а)},
Дмитрий Дмитриевич Кенчадзе^{б)},
Кирилл Александрович Алексеев^{б)}

^{а)} Национальный исследовательский университет «Высшая школа экономики», г. Москва, Россия;

^{б)} Федеральная служба государственной статистики, г. Москва, Россия

Процедура комплексирования частично известного вектора квартальных данных рассматривается в статье как способ идентификации неизвестных компонент этого вектора на основе структурной информации, содержащейся в его известных компонентах и в заранее выбранной эталонной макроэкономической матрице. Иными словами, рассматриваемая задача комплексирования вектора квартальных данных сводится к «достраиванию» неизвестной его части по известной на основе косвенной априорной информации.

Статью открывает формальная постановка общей задачи комплексирования вектора вместе с ее экономико-статистической интерпретацией для случая согласования квартальных счетов производства. Получено общее аналитическое решение задачи комплексирования вектора строчных сумм на основе блочного разбиения эталонной матрицы с неотрицательными элементами в виде линейной модели. Показано, как применять полученное решение для комплексирования вектора выпуска продуктов в отчетном квартале с использованием эталонной матрицы, ассоциированной с соответствующим кварталом предыдущего года, и как эвристически оценивать погрешность расчетов для отчетного квартала, анализируя отклонение расчетного вектора от гомотетического луча, определяемого эталонной матрицей. Предложено обобщение аналитического решения задачи комплексирования для совместной корректировки квартальных итогов производства в отчетном году с целью приведения их в точное соответствие с годовыми данными. Особое внимание уделено эвристической оценке погрешности расчетов для отчетного года на основе анализа различия между системой расчетных квартальных векторов и гомотетией их суммы (вектора-столбца окаймляющих итогов эталонной годовой матрицы). В заключение даны рекомендации по повышению надежности результатов комплексирования квартальных векторов в практических ситуациях.

Ключевые слова: задача комплексирования вектора, эталонная матрица, блочное разбиение матрицы, комплексирование в отчетном квартале, комплексирование в отчетном году, оценки погрешности расчетов.

JEL: E01, C61, C65.

doi: <https://doi.org/10.34023/2313-6383-2022-29-3-68-77>.

Для цитирования: Моторин В.И., Кенчадзе Д.Д., Алексеев К.А. Модели комплексирования квартальных данных на основе структурных характеристик эталонной матрицы. Вопросы статистики. 2022;29(3):68–77.

Models for Completing the Quarterly Data by Using the Structural Characteristics of a Reference Matrix

Vladimir I. Motorin^{a)},
Dmitry D. Kenchadze^{b)},
Kirill A. Alekseev^{b)}

^{a)} National Research University Higher School of Economics (HSE University), Moscow, Russia;

^{b)} Federal State Statistics Service, Moscow, Russia

Completing a partially known vector of quarterly data is considered in the article as a way to identify unknown components of this vector based on structural information contained in its known components as well as in the preselected reference macroeconomic matrix. In other words, the problem of completing a quarterly vector is reduced to «upbuilding» its unknown part according to the known one based on indirect a priori information.

The article opens with a formal statement of the general problem of completing vector data together with its economic and statistical interpretation as applying to reconciling the quarterly production accounts. A general analytical solution to the problem of completing the vector of row sums is obtained on the basis of block partitioning of a reference matrix with nonnegative elements in the form of a linear model.

It is shown how to use the obtained solution for completing the product output vector in the reporting quarter using the reference matrix associated with the corresponding quarter of the previous year, and how to heuristically estimate the calculation imprecision rate for the reporting quarter by analyzing the deviation of the calculated vector from the homothetic ray determined by the reference matrix. Further generalization of the analytical solution of the completing problem for the joint adjustment of quarterly production results in the reporting year in order to bring them into exact correspondence with the annual data is proposed. Particular attention is paid to heuristic estimating the imprecision rate in calculations for the reporting year based on the analysis of the difference between the system of calculated quarterly vectors and a homothety of their sum (i. e., column vector of the row margin totals for the reference annual matrix). In conclusion, recommendations to improve reliability of the completion results for quarterly vectors in practical situations are given.

Keywords: completion of a vector, reference matrix, matrix block partitioning, completing data in reporting quarter, completing data in reporting year, estimates of calculation imprecision rate.

JEL: E01, C61, C65.

doi: <https://doi.org/10.34023/2313-6383-2022-29-3-68-77>.

For citation: Motorin V.I., Kenchadze D.D., Alekseev K.A. Models for Completing the Quarterly Data by Using the Structural Characteristics of a Reference Matrix. *Voprosy Statistiki*. 2022;29(3):68–77. (In Russ.)

Универсальная задача моделирования векторных и матричных макроэкономических структур при построении квартальных национальных счетов *в режиме отчетного квартала* сводится к построению априори неизвестной прямоугольной квартальной матрицы по частично или полностью известному вектору-столбцу ее окаймляющих итогов и с использованием других структурных данных за соответствующий квартал предыдущего года. В свою очередь, универсальная задача моделирования векторных и матричных макроэкономических структур при разработке квартальных счетов *в режиме отчетного года* заключается в построении комплекта априори неизвестных прямоугольных квартальных матриц, сумма которых равна заданной годовой матрице, по набору частично или полностью известных векторов-столбцов окаймляющих итогов этих квартальных матриц и годовой матрице.

Первичные оценки макроэкономических показателей ввиду обычного для статистической практики недостатка исходных данных формируют, как правило, лишь часть компонент квартальных векторов с элементами, исчисляемыми в разрезе продуктов. В связи с этим при построении квартальных счетов возникает необходимость оценивания остальных (неизвестных) компонент квартальных векторов с использованием тех или иных косвенных макроэкономических данных и математических методов. Таким образом, при моделировании векторных и матричных макроэкономических структур для построения квартальных счетов требуется реализовать процедуру комплексирования частично известных квартальных векторов на основе структурных характеристик априори заданной эталонной матрицы,

а также (в режиме отчетного года) обеспечить согласование построенных квартальных векторов с годовыми данными.

Настоящая статья посвящена разработке оригинальных алгоритмов и процедур указанного комплексирования и согласования векторов. Разработанные алгоритмы носят универсальный характер, но тем не менее для удобства изложения и облегчения восприятия материала текст статьи подготовлен на лексическом примере квартальных и годовых оценок выпуска продуктов и отраслей.

Формальная постановка общей задачи комплексирования вектора

Пусть имеется прямоугольная матрица V размерности $N \times M$ с неотрицательными элементами и N -мерным вектором-столбцом строчных сумм $r = Ve_M$, где e_M – суммирующий вектор (англ. *summation vector*, вектор с единичными компонентами) размерности $M \times 1$. Представим вектор r в виде конкатенации подвектора-столбца r_1 размерности $L \times 1$, где $L < N$, и подвектора-столбца r_2 размерности $(N - L) \times 1$, то есть положим

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}.$$

Кроме того, пусть задан некоторый вектор $\tilde{r}_1 \neq r_1$ размерности $L \times 1$ как часть N -мерного вектора-столбца

$$\tilde{r} = \begin{bmatrix} \tilde{r}_1 \\ \tilde{r}_2 \end{bmatrix} \neq r \quad (1)$$

с неизвестным подвектором-столбцом $\tilde{r}_2$ размерности $(N - L) \times 1$. Вектор (1) здесь рассматрива-

ется как вектор строчных сумм элементов некоторой неизвестной матрицы $\tilde{V}$, часть компонент которого задана априори.

Общая задача комплексирования частично известного вектора строчных сумм $\tilde{r}$ заключается в использовании экзогенного подвектора $\tilde{r}_1$ и структурной информации, содержащейся в исходной (эталонной) матрице V , для идентификации компонент неизвестного подвектора $\tilde{r}_2$. Таким образом, под процедурой комплексирования вектора здесь понимается «достраивание» неизвестной его части по известному подвектору на основе косвенных априорных данных.

Подобная задача известна в теории машинного обучения (англ. *machine learning*) под собирательным термином «завершение матрицы» (англ. *matrix completion*) и связана с восстановлением неизвестных элементов частично наблюдаемой матрицы (см., например, [1]). Следует подчеркнуть, что доминирующим инструментальным подходом в этой научной области является использование критерия минимизации ранга «завершенной» матрицы, соответствующие элементы которой в точности совпадают с известными априори [2]. Операциональность подхода опирается на доказательство того факта, что при выборочном характере наблюдений и их достаточном количестве задача минимизации ранга искомой матрицы с высокой вероятностью имеет единственное решение [3]. Однако вполне очевидно, что столь формальный критерий не применим к таким взаимосвязанным объектам макроэкономической статистики, как квартальные национальные счета.

Матрица выпуска и экономико-статистическая интерпретация задачи

В мировой статистической практике общая задача комплексирования вектора в сформулированной выше постановке ранее не рассматривалась. Тем не менее она представляется вполне актуальной и полезной при разработке и балансировке счетов производства (квартальных и годовых) в условиях, когда своевременное построение полной системы предварительных оценок суммарных объемов выпуска продуктов по кварталам данного года на основе первичных статистических данных (англ. *basic statistics, basic data*) сопряжено со значительными информационными трудностями [4] и некоторые оценки выпуска продукции оказы-

ваются настолько ненадежными, что их лучше исключить из дальнейших расчетов (см. [5], раздел «Неполнота системы экзогенных предварительных оценок»). При этом V здесь интерпретируется как прямоугольная матрица выпуска продукции в стоимостном выражении по N видам экономической деятельности (продуктам) и по M отраслям экономики (англ. *production matrix*), а r — как вектор суммарных объемов выпуска каждого продукта во всех отраслях (англ. *product marginal totals*).

Применительно к счетам производства процедура комплексирования вектора оказывается полезной в двух практических ситуациях.

Первая из них связана с построением квартальных счетов производства в режиме отчетного квартала (за I, II и III кварталы текущего года) в условиях, когда годовая матрица выпуска продукции еще не известна. В этой ситуации роль эталона V играет матрица выпуска продукции в соответствующем квартале предыдущего года, а подвектор $\tilde{r}_1$ определяется на основе первичных статистических данных (и логических схем их обработки, представленных на Fig. 4.1 и Fig. 5.1 в авторитетных международных руководствах [6 и 7] соответственно) или путем построения актуальных и относительно надежных квартальных оценок, например на базе информации об объемах отгруженных товаров собственного производства, выполненных работ и услуг собственными силами по отдельным видам экономической деятельности с помощью усовершенствованного пропорционального метода экстраполяции Дентона (англ. *enhanced proportional Denton extrapolation method*) [8, par. 6.33, 6.34] или пропорционального метода бенчмаркинга Шолетта — Дагум с авторегрессионной ошибкой первого порядка (англ. *proportional Cholette — Dagum benchmarking method with first-order autoregressive error*) [9; 10, par. 6.51–6.61].

Другая ситуация возникает при корректировке квартальных счетов производства в режиме отчетного года с целью приведения их в точное соответствие с годовыми данными о выпуске продукции. Здесь в качестве эталона V естественно выбрать годовую матрицу выпуска продукции, а набор квартальных подвекторов $\tilde{r}_1$, компоненты которых согласованы с годовыми окаймляющими итогами r_1 , определяется на основе построенных ранее фрагментов квартальных счетов для отчетного года с помощью специальных методов — например, с использованием оригинальных результатов, изложенных в статье [5].

Решение общей задачи комплексирования вектора

Для решения поставленной задачи требуется установить формальную связь между векторными фрагментами r_1 и r_2 окаймляющих итогов заданной эталонной матрицы V . Постулируя эту связь в линейной форме, положим $r_2 = F_V r_1$, где F_V — неизвестная прямоугольная матрица размерности $(N-L) \times L$ с элементами, зависящими от данных о структуре матрицы V .

В качестве структурных характеристик матрицы V естественно рассмотреть две $N \times M$ матрицы ее вертикальной и горизонтальной структуры:

$$G = V \langle e'_N V \rangle^{-1} = V \hat{c}^{-1} \text{ и } H = \langle V e_M \rangle^{-1} V = \hat{r}^{-1} V, \quad (2)$$

где $c = V e'_N$ и $r = V e_M$ — векторы-столбцы столбцовых и строчных сумм элементов V размерности $M \times 1$ и $N \times 1$ соответственно, надстрочным символом «штрих» обозначена операция транспонирования матриц (векторов), а заключение символа вектора в угловые скобки « $\langle \cdot \rangle$ » (или знак « $\wedge$ » над ним) обозначает его преобразование в диагональную матрицу соответствующего порядка с элементами этого вектора на главной диагонали [11, р. 697]. Здесь неявно предполагается, что данная матрица V с неотрицательными элементами не содержит нулевых строк и/или столбцов.

Из определений (2) следует, что матрицу V можно представить как в виде $V = G \hat{c}$ так и в альтернативном виде $V = \hat{r} H$, откуда

$$r = V e_M = G \hat{c} e_M = G c, \quad c = V e'_N = H \hat{r} e'_N = H r. \quad (3)$$

После транспонирования левой и правой частей второй формулы (3) на основе дихотомического разбиения матрицы H имеем:

$$c' = r' H = (r'_1, r'_2) \begin{pmatrix} E'_N H \\ E'_N H \end{pmatrix} = r'_1 E'_N H + r'_2 E'_N H, \quad (4)$$

где E'_N — прямоугольная матрица размерности $L \times N$, образованная верхними L строками единичной матрицы E_N порядка N , а E'_N — прямоугольная матрица размерности $(N-L) \times N$, образованная нижними $N-L$ строками той же единичной матрицы. Таким образом, матрица $E'_N H$ охватывает верхние L строк матрицы H , а матрица $E'_N H$ состоит из ее нижних $N-L$ строк.

Подставив транспонированное соотношение (4) в первую формулу (3), получим уравнение связи векторов r , r_1 и r_2 в линейной форме:

$$r = G c = G H E'_N r_1 + G H E'_N r_2.$$

Умножая обе его части на матрицу E_N слева и учитывая то очевидное обстоятельство, что $E_N r = r_2$, имеем:

$$r_2 = E_N G H E'_N r_1 + E_N G H E'_N r_2.$$

Перегруппировав члены и разрешив последнее уравнение относительно r_2 , получим следующую линейную зависимость вектора r_2 от r_1 :

$$r_2 = (E_{N-L} - E_N G H E'_N)^{-1} E_N G H E'_N r_1, \quad (5)$$

где квадратная матрица в круглых скобках имеет порядок $N-L$.

Здесь важно подчеркнуть, что при $r' = (r'_1, r'_2) = e'_M V'$ уравнение (5) представляет собой тождество, справедливость которого, как показано выше, вытекает из равенств (2), определяющих структурные матрицы G и H .

Комплексирование вектора выпуска продуктов в отчетном квартале

Установленная линейная связь $r_2 = F_V r_1$ между векторными фрагментами r_2 и r_1 окаймляющих итогов заданной матрицы V может быть использована при проведении расчетов в процессе построения квартальных счетов производства в режиме отчетного квартала. Как отмечалось выше, в этом случае вектор r_1 представляет собой актуальный набор относительно надежных квартальных данных о выпуске продуктов в отчетном квартале текущего года $\tilde{r}_1$, а в качестве эталона V выступает матрица выпуска продукции в соответствующем квартале предыдущего года.

Расчет вектора выпуска продуктов в отчетном квартале текущего года $\tilde{r}$ предлагается производить по формуле

$$\tilde{r} = \begin{bmatrix} \tilde{r}_1 \\ \tilde{r}_2 \end{bmatrix} = \begin{bmatrix} \tilde{r}_1 \\ F_V \tilde{r}_1 \end{bmatrix}, \quad (6)$$

где, как следует из уравнения связи (5), $F_V = (E_{N-L} - E_N G H E'_N)^{-1} E_N G H E'_N$ — прямоугольная матрица размерности $(N-L) \times L$. Однако

здесь необходимо подчеркнуть, что правомерность предпринятой в (6) замены «исторического» тождества $r_2 = F_V r_1$, справедливого в рамках некоторого интервала времени в прошлом, линейной моделью $\tilde{r}_2 = F_V \tilde{r}_1$, связывающей актуальные макроэкономические данные, вообще говоря, не является очевидной и не может быть доказана на основе каких бы то ни было формальных утверждений.

Эвристическая оценка погрешности расчетов для отчетного квартала

Обоснование возможности практического использования линейной модели $\tilde{r}_2 = F_V \tilde{r}_1$ для комплексирования вектора выпуска продуктов в отчетном квартале текущего года опирается на следующие соображения. Поочередно подставив одно из равенств (3) в другое, получим два характерных уравнения

$$r = GH'r \quad \text{и} \quad c = H'Gc \quad (7)$$

относительно векторов r и c соответственно.

Как следует из (7), N -мерный вектор-столбец $r = Ve_M$ является собственным вектором квадратной матрицы GH' порядка N , ассоциированным с единичным собственным значением этой матрицы. Аналогично: квадратная матрица $H'G$ порядка M также имеет единичное собственное число, которому соответствует M -мерный собственный вектор-столбец $c = Ve_N$. Итак, матрицы GH' и $H'G$ однозначно определяют окаймляющие итоги заданной матрицы V и поэтому вполне обоснованно могут рассматриваться как важные структурные характеристики этой матрицы [здесь следует подчеркнуть, что GH' фигурирует в (5) в качестве единственного носителя информации о заданной матрице V]. Отметим также, что структурные матрицы выпуска G и H играют существенную роль в математическом инструментарии анализа «затраты-выпуск» [11 и 12].

Разумеется, при $\tilde{V} \neq V$ и $\tilde{r}_1 = E_N \tilde{V} e_M \neq r_1$ вектор $\tilde{r}$, рассчитанный по предложенной формуле (6), вообще говоря, не удовлетворяет уравнению $\tilde{r} = GH'\tilde{r}$, то есть не является собственным вектором матрицы GH' . Однако если $\tilde{V} = kV$, где k — произвольный ненулевой скаляр, то структурные характеристики матрицы выпуска остаются неизменными, так как

$$\tilde{G} = \tilde{V} \langle e'_N \tilde{V} \rangle^{-1} = kV \langle ke'_N V \rangle^{-1} = G$$

и

$$\tilde{H} = \langle \tilde{V} e_M \rangle^{-1} \tilde{V} = \langle kV e_M \rangle^{-1} kV = H.$$

Далее, при подстановке $\tilde{r}_1 = kr_1$ в (6) имеем $\tilde{r}_2 = F_V \tilde{r}_1 = kF_V r_1 = kr_2$, откуда $\tilde{r} = kr$. Таким образом, использование формулы (6) в задаче комплексирования частично известного квартального вектора приводит к получению точного решения в ситуации, когда экзогенные данные формируются на основе гомотетического (пропорционального) преобразования исходных данных (более подробные сведения о гомотетическом тестировании методов актуализации векторно-матричных структур можно найти в [13]).

Вместе с тем ясно, что в общем случае экзогенный подвектор $\tilde{r}_1$ не является гомотетией r_1 , а рассчитанный по формуле (6) вектор $\tilde{r}$ не удовлетворяет уравнению $\tilde{r} = GH'\tilde{r}$. Как следствие, $\tilde{r}$ может рассматриваться как вектор-столбец строчных сумм неизвестной матрицы $\tilde{V}$, структура которой близка к структуре исходной эталонной матрицы V , лишь с некоторой погрешностью, характеризующей надежность результатов выполненных расчетов. Для того чтобы получить представление о надежности решения (6) задачи комплексирования частично известного вектора выпуска продуктов в отчетном квартале, требуется провести количественную оценку различия между вектором $\tilde{r}$ и N -мерным вектором-столбцом r строчных сумм эталонной матрицы.

Здесь важно сразу же подчеркнуть, что использование в этих целях разности $\tilde{r} - r$ не является логически обоснованным, поскольку первое из уравнений (7) определяет вектор r с точностью до произвольного скалярного множителя. Следовательно, на самом деле необходимо оценить различие не между двумя векторами $\tilde{r}$ и r , а между вектором $\tilde{r}$ и гомотетическим лучом kr . Это можно сделать как на основе тригонометрического подхода, так и в рамках геометрических построений.

Величина угла между $\tilde{r}$ и лучом kr (в радианах) определяется по известной формуле:

$$\beta_{r,\tilde{r}} = \arccos \left[\frac{r'\tilde{r}}{(r'r)^{1/2} \cdot (\tilde{r}'\tilde{r})^{1/2}} \right]. \quad (8)$$

Значения $\beta_{r,\tilde{r}}$, близкие к нулю, свидетельствуют о высокой надежности решения (6), тогда как увеличение $\beta_{r,\tilde{r}}$ от 0 до $\pi/2$ постепенно выводит это решение за грань практической приемлемости. При проведении расчетов с величиной угла (8) удобно связать безразмерный нормированный показатель

$$\delta_{\perp} = 2\beta_{r,\tilde{r}}/\pi, \quad (9)$$

принимающий значения в диапазоне от 0 до 1 и выражающий $\beta_{r,\tilde{r}}$ в долях величины прямого угла $\pi/2$.

Другой — геометрический — способ оценки различия между вектором $\tilde{r}$ и лучом kr опирается на операцию ортогонального проектирования вектора $\tilde{r}$ на гомотетический луч. Условие ортогональности векторов kr и $\tilde{r} - kr$, записанное в виде равенства нулю их скалярного произведения $r'(\tilde{r} - kr) = 0$, позволяет определить относительную координату точки проекции на гомотетическом луче $k_{\perp} = r'\tilde{r}/r'r$. Построенный в результате вектор $d = \tilde{r} - k_{\perp}r$ соответствует минимальному отклонению вектора $\tilde{r}$ от точек гомотетического луча, а его модуль может служить метрической оценкой различия между ними. Таким образом, безразмерный нормированный показатель

$$\delta_{\perp} = \left[\frac{(\tilde{r}' - k_{\perp}r')(\tilde{r} - k_{\perp}r)}{\tilde{r}'\tilde{r}} \right]^{1/2}, \quad k_{\perp} = \frac{r'\tilde{r}}{r'r}, \quad (10)$$

принимает значения в диапазоне от 0 до 1 и в определенном смысле характеризует степень надежности решения (6) задачи комплексирования частично известного вектора выпуска продуктов в отчетном квартале текущего года, как и угловая мера (9). Из геометрических построений следует, что показатель (10) связан с величиной угла (8) соотношением $\delta_{\perp} = \sin \beta_{r,\tilde{r}}$.

Корректировка квартальных векторов выпуска продуктов в отчетном году

Линейную модель $r_2 = F_V r_1$, связывающую векторные фрагменты r_2 и r_1 окаймляющих итогов заданной матрицы V , нетрудно обобщить для корректировки квартальных счетов производства в режиме отчетного года с целью приведения их в точное соответствие с годовыми данными о выпуске продукции. Выше отмечалось, что в этой за-

даче в качестве эталонной матрицы V фигурирует годовая матрица выпуска продукции в отчетном году, а исходная квартальная информация представлена построенными ранее квартальными сче- тами для отчетного года, обычно не согласованными с годовыми данными как количественно, так и структурно.

Логика предлагаемой процедуры корректировки квартальных векторов выпуска продуктов в отчетном году заключается в следующем. Необходимость обеспечить точное соответствие между квартальными и годовыми данными о выпуске продукции в практической статистической работе порождает задачу квартальной декомпозиции продуктовых окаймляющих итогов годовой матрицы выпуска с привлечением некоторого количества более или менее надежных экзогенных данных в квартальном разрезе, селективируемых из построенных ранее квартальных счетов для отчетного года с использованием экспертных оценок и различных специальных методов [6—8 и 10]. В результате указанной селекции на практике обычно оказывается возможным сформировать набор предварительных оценок квартальных выпусков всех или (гораздо чаще) нескольких продуктов — оценок относительно надежных, но, к сожалению, не согласованных с годовыми отчетными данными.

Применение к этому набору алгоритмов согласования, разработанных в статье [5], позволяет построить информационное «ядро» для последующей компиляции квартальных счетов производства для отчетного года, сбалансированных с годовыми данными. Формально оно может быть представлено в виде прямоугольной матрицы $\tilde{R}_1$ размерности $L \times T$, составленной из вычисленных (в процессе упомянутого выше согласования) квартальных векторов-столбцов выпуска продуктов $\tilde{r}_1'$ ($t = 1 \div T$) размерности $L \times 1$, где T — число периодов времени (кварталов) в отчетном году, а символ « $\div$ » между нижней и верхней границами диапазона изменения индекса t означает, что этот индекс последовательно пробегает все целочисленные значения в указанном диапазоне. Заметим, что порядок следования L продуктов в экзогенно сформированной матрице $\tilde{R}_1$ должен быть в точности таким же, как в верхней части годовой матрицы выпуска продукции в отчетном году V ; в противном случае строки одной из матриц следует переставить надлежащим образом.

Пусть, как и ранее, вектор r_1 размерности $L \times 1$ есть подвектор N -мерного столбца строчных сумм $r = Ve_M$ эталонной матрицы V , а вектор r_2 размерности $(N-L) \times 1$ является дихотомическим дополнением r_1 до r . В этих обозначениях соответствие экзогенной матрицы $\tilde{R}_1$ годовым данным о выпуске продуктов означает, что $\tilde{R}_1 e_T = r_1$.

Далее, пусть неизвестная матрица $\tilde{R}_2$ размерности $(N-L) \times T$ является дихотомическим дополнением $\tilde{R}_1$ до матрицы квартальных продуктовых итогов $\tilde{R}$ размерности $N \times T$. Ясно, что она должна удовлетворять балансовому условию $\tilde{R}_2 e_T = r_2$. Выше показано, что векторные фрагменты r_2 и r_1 окаймляющих итогов эталонной матрицы V связаны матричным соотношением (5). Подставив уравнения $\tilde{R}_2 e_T = r_2$ и $\tilde{R}_1 e_T = r_1$ в левую и правую части (5) соответственно, имеем

$$\tilde{R}_2 e_T = (E_{N-L} - E_N GH'E_N')^{-1} E_N GH'E_N' \tilde{R}_1 e_T,$$

или

$$\sum_{t=1}^T \tilde{r}_2^t = (E_{N-L} - E_N GH'E_N')^{-1} E_N GH'E_N' \sum_{t=1}^T \tilde{r}_1^t. \quad (11)$$

Здесь необходимо отметить, что с формальной точки зрения (11) представляет собой систему $N-L$ линейных уравнений относительно $(N-L)T > N-L$ неизвестных компонент T векторов $\tilde{r}_2^t$ размерности $(N-L) \times 1$, $t = 1 \div T$, очевидно имеющую бесконечное множество решений и, как следствие, не составляющую самостоятельную практической ценности.

Простейший прием превращения системы уравнений (11) в однозначно идентифицируемую модель заключается в расщеплении (11) на T групп по $N-L$ линейных уравнений, в каждой из которых фигурирует только один подлежащий идентификации неизвестный вектор $\tilde{r}_2^t$ размерности $(N-L) \times 1$, $t = 1 \div T$. Выполнив процедуру расщепления, получим

$$\tilde{r}_2^t = (E_{N-L} - E_N GH'E_N')^{-1} E_N GH'E_N' \tilde{r}_1^t, \quad t = 1 \div T. \quad (12)$$

Важно подчеркнуть, что системы линейных уравнений (11) и (12) не являются эквивалентными, хотя единственное решение системы (12), конечно же, представляет собой одно из допустимых решений системы (11).

Таким образом, расчет квартальных векторов выпуска продуктов в отчетном году $\tilde{r}^t$ предлагается производить по формуле

$$\tilde{r}^t = \begin{bmatrix} \tilde{r}_1^t \\ \tilde{r}_2^t \end{bmatrix} = \begin{bmatrix} \tilde{r}_1^t \\ F_V \tilde{r}_1^t \end{bmatrix}, \quad t = 1 \div T, \quad (13)$$

где, как и в (5), $F_V = (E_{N-L} - E_N GH'E_N')^{-1} E_N GH'E_N'$ — прямоугольная матрица размерности $(N-L) \times L$.

Заметим, что если сумма всех T квартальных векторов-столбцов выпуска продуктов $\tilde{r}_1^t$ равна годовому вектору r_1 , то сумма вычисленных по формуле (13) векторных оценок $\tilde{r}^t$ в точности равняется годовому вектору выпуска продуктов r , так как $r_2 = F_V r_1$. Следовательно, формула (13) обеспечивает требуемую сбалансированность системы квартальных оценок выпуска продуктов с годовыми данными.

Эвристическая оценка погрешности расчетов для отчетного года

Нетрудно видеть, что использование формулы (13) в задаче корректировки квартальных векторов выпуска продуктов в режиме отчетного года обеспечивает получение точного решения в ситуации, когда столбцы экзогенной матрицы $\tilde{R}_1$ формируются на основе гомотетического (пропорционального) преобразования годового вектора выпуска продуктов.

Однако в общем случае каждый из T столбцов матрицы $\tilde{R}_1$ не выражается гомотетией вектора r_1 , а рассчитанные по формуле (13) векторы $\tilde{r}^t$ не являются собственными векторами характеристической матрицы GH' , то есть не удовлетворяют уравнению $\tilde{r}^t = GH' \tilde{r}^t$. Поэтому каждый из столбцов $\tilde{r}^t$ может рассматриваться как вектор-столбец строчных сумм неизвестной квартальной матрицы $\tilde{V}^t$, структура которой близка к структуре эталонной годовой матрицы V , лишь с некоторой погрешностью, характеризующей надежность результатов расчетов по формуле (13) для t -го периода времени (квартала).

Для того чтобы получить представление о надежности решения (13) задачи корректировки квартальных векторов выпуска продуктов в режиме отчетного года, требуется провести количественную оценку различия между системой T квартальных векторов $\tilde{r}^t$ ($t = 1 \div T$) и гомотетией их суммы — вектора-столбца окаймляющих ито-

гов $r = Ve_M$ эталонной (годовой) матрицы выпуска V . Как и ранее при анализе задачи комплексирования частично известного квартального вектора выпуска в режиме отчетного квартала, это можно сделать не только в рамках тригонометрического подхода на базе угловой меры (8), но и на основе геометрических построений с использованием показателя (10).

Угловую меру надежности решения (13) логично построить путем арифметического усреднения T углов между каждым из построенных квартальных векторов $\tilde{r}^t$ и годовым вектором выпуска $r = Ve_M$:

$$\bar{\delta}_\angle = \frac{2}{\pi T} \sum_{t=1}^T \beta(r, \tilde{r}^t), \quad (14)$$

где $\beta(r, \tilde{r}^t)$ – угол между r и $\tilde{r}^t$, вычисляемый по аналогии с (8). Безразмерный показатель (14) изменяется в диапазоне от 0 до 1 и выражает среднее значение углов между $\tilde{r}^t$ и r в долях величины прямого угла $\pi/2$.

Другой способ оценки надежности решения (3) связан с операцией ортогонального проектирования векторов $\tilde{r}^t$ на гомотетический луч kr . Из условия ортогональности векторов kr и $\tilde{r}^t - kr$, записанного в виде равенства нулю их скалярного произведения $r'(\tilde{r}^t - kr) = 0$, можно определить относительную координату t -й точки проекции на гомотетическом луче $k_\perp^t = r' \tilde{r}^t / r'r$, $t = 1 \div T$. Каждый из T полученных векторов $d^t = \tilde{r}^t - k_\perp^t r$ соответствует минимальному отклонению вектора $\tilde{r}^t$ от точек гомотетического луча, а обобщенной метрической оценкой различия между системой T квартальных векторов $\tilde{r}^t$ и лучом kr может служить среднее значение модулей векторов d^t ($t = 1 \div T$). В связи с этим нетрудно видеть, что безразмерный нормированный показатель

$$\bar{\delta}_\perp = \left[\sum_{t=1}^T (\tilde{r}^t - k_\perp^t r)' (\tilde{r}^t - k_\perp^t r) / \sum_{t=1}^T (\tilde{r}^t)' \tilde{r}^t \right]^{1/2},$$

$$k_\perp^t = \frac{r' \tilde{r}^t}{r'r}, \quad (15)$$

принимает значения в диапазоне от 0 до 1 и в известном смысле характеризует степень надежности решения (13) задачи корректировки квартальных векторов выпуска продуктов в отчетном году, как и угловая мера (14). В отличие от предыдущего случая анализа данных для отчетного

квартала по формулам (9) и (10), измерители (14) и (15) не связаны каким-либо функциональным соотношением и поэтому являются взаимодополняющими.

Заключительные замечания и практические соображения

Важный в практическом отношении вопрос об определении верхних границ приемлемости решений (6) и (13) в терминах показателей (9), (10) и (14), (15) весьма затруднительно разрешить на основе теоретических конструкций из области математической статистики и эконометрики. Представляется, что более или менее исчерпывающий ответ на него можно получить только эмпирически, путем проведения массированных расчетов на базе фактической статистической информации.

По результатам выполненной экспериментальной апробации разработанных методов комплексирования квартальных векторов выпуска продуктов на основе специально подготовленных массивов искусственных данных можно утверждать, что при наблюдаемых уровнях значений индикаторов (9), (10) и (14), (15) менее 10% получаемые решения демонстрируют достаточно высокую надежность, а при уровне от 10 до 20% их допустимо считать условно надежными.

Значения индикаторов на уровнях выше 20% свидетельствуют о весьма значительных структурных сдвигах в экзогенной информации по сравнению с эталонными данными, не позволяющих рассчитывать на получение сколько-нибудь надежных решений. Вместе с тем исходя из накопленного опыта обработки макроэкономических данных есть веские основания полагать, что вероятность столкнуться с подобными явлениями на макроуровне в реальной статистической практике чрезвычайно мала.

Тем не менее если расчетные значения индикаторов (9), (10) и (14), (15) представляются неудовлетворительными, можно попытаться дополнить информационное «ядро» проводимых вычислений экзогенными квартальными оценками выпуска того или иного продукта, данные о квартальных выпусках которого до сих пор фигурировали в расчетах как эндогенные (разумеется, если первичные статистические данные для подготовки такого дополнения имеются).

Ясно, что при этом в задаче комплексирования квартального вектора выпуска продуктов в режиме отчетного квартала размерность известного вектора $\tilde{r}_1$ увеличивается до $(L+1) \times 1$, а размерность неизвестного вектора $\tilde{r}_2$ уменьшается до $(N-L-1) \times 1$. Аналогично в задаче корректировки квартальных векторов выпуска продуктов в режиме отчетного года размерность известной прямоугольной матрицы $\tilde{R}_1$ возрастает до $(L+1) \times T$, а неизвестная матрица $\tilde{R}_2$ теряет одну строку и эволюционирует до размерности $(N-L-1) \times T$. Представляется вполне вероятным, что в новом цикле расчетов с измененными исходными данными значения индикаторов (9), (10) и (14), (15) уменьшатся до удовлетворительного уровня.

В заключение следует заметить, что изложенные выше методы комплексирования квартальных векторов выпуска продуктов на основе структурных характеристик эталонной матрицы выпуска без каких бы то ни было препятствий могут быть использованы в задачах комплексирования квартальных векторов выпуска отраслей. Для такого применения разработанных методов и проведения расчетов по формулам (6) и (13) необходимо подготовить соответствующий набор экзогенных квартальных оценок в отраслевом разрезе и транспонировать эталонную матрицу выпуска.

Литература

1. Recht B. A Simpler Approach to Matrix Completion // Journal of Machine Learning Research. 2011. Vol. 12. P. 3413–3430.
2. Liu Z., Hu Z., Nie F. Matrix Completion and Vector Completion via Robust Subspace Learning // Neurocomputing. 2018. Vol. 306. P. 171–181. doi: <https://doi.org/10.1016/j.neucom.2018.04.032>.
3. Candès E.J., Recht B. Exact Matrix Completion via Convex Optimization // Foundations of Computational Mathematics. 2009. Vol. 9. Iss. 6. P. 717–772. doi: <https://doi.org/10.1007/s10208-009-9045-5>.
4. Landefeld J.S., Seskin E.P., Fraumeni B.M. Taking the Pulse of the Economy: Measuring GDP // Journal of Economic Perspectives. 2008. Vol. 22. No. 2. P. 193–216. doi: <https://doi.org/10.1257/jep.22.2.193>.
5. Моторин В.И., Кенчадзе Д.Д. Методы согласования квартальных оценок выпуска продуктов и отраслей с годовыми данными о выпуске продукции // Вопросы статистики. 2017. № 10. С. 3–12.
6. European Commission (Eurostat). Handbook on Quarterly National Accounts. Luxembourg: Publications Office of the European Union, 2013. 312 p. doi: <https://doi.org/10.2785/46080>.
7. European Commission (Eurostat). Handbook on Quarterly National Accounts: Methods and Nomenclatures. Luxembourg: Office for Official Publications of the European Communities, 1999. 436 p.
8. Блум Э.М., Диппелсман Р.Дж., Меле Н.Э. Руководство по квартальным национальным счетам: концепции, источники данных и составление: пер. с англ. Вашингтон, округ Колумбия: МВФ, 2001. 222 с.
9. Dagum E.B., Cholette P.A. Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series. Lecture Notes in Statistics # 186. New York: Springer Science+Business Media, 2006. 409 p.
10. Руководство по квартальным национальным счетам: издание 2017 года: пер. с англ. Вашингтон, округ Колумбия: МВФ, 2018. 315 с.
11. Miller R.E., Blair P.D. Input-Output Analysis: Foundations and Extensions: 2nd ed. N.Y.: Cambridge University Press, 2009. 782 p. doi: <https://doi.org/10.1017/CBO9780511626982>.
12. Motorin V. The Linear Matrix-Valued Cost Functions as a Source of Leontief and Ghosh Model // Artha Vijnana: Journal of the Gokhale Institute of Politics and Economics. 2017. Vol. 59. No. 4. P. 291–301. doi: <https://doi.org/10.21648/arthavij/2017/v59/i4/170795>.
13. Motorin V. Enhancing the Distance Minimization Methods of Matrix Updating Within a Homothetic Paradigm // Journal of Economic Structures. 2017. Vol. 6. Article number: 36. P. 1–22. doi: <https://doi.org/10.1186/s40008-017-0094-7>.

Информация об авторах

Владимир Ильич Моторин — канд. экон. наук, старший научный сотрудник, главный эксперт Центра экономических измерений и статистики, Департамент статистики и анализа данных, Национальный исследовательский университет «Высшая школа экономики». 109028, г. Москва, Покровский бульвар, д. 11. E-mail: motoriny@yandex.ru. ORCID: <https://orcid.org/0000-0003-0924-8275>.

Дмитрий Дмитриевич Кенчадзе — заместитель руководителя, Федеральная служба государственной статистики. 107450, г. Москва, ул. Мясницкая, д. 39, стр. 1. E-mail: Ken Chadze@gks.ru.

Кирилл Александрович Алексеев — заместитель начальника управления национальных счетов, Федеральная служба государственной статистики. 107450, г. Москва, ул. Мясницкая, д. 39, стр. 1. E-mail: AlekseevKA@gks.ru.

Финансирование

Исследование осуществлено в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ).

References

1. **Recht B.** A Simpler Approach to Matrix Completion. *Journal of Machine Learning Research*. 2011;12(104):3413–3430.
2. **Liu Z., Hu Z., Nie F.** Matrix Completion and Vector Completion via Robust Subspace Learning. *Neurocomputing*. 2018;306:171–181. Available from: <https://doi.org/10.1016/j.neucom.2018.04.032>.
3. **Candès E.J., Recht B.** Exact Matrix Completion via Convex Optimization. *Foundations of Computational Mathematics*. 2009;9(6):717–772. Available from: <https://doi.org/10.1007/s10208-009-9045-5>.
4. **Landefeld J.S., Seskin E.P., Fraumeni B.M.** Taking the Pulse of the Economy: Measuring GDP. *Journal of Economic Perspectives*. 2008;22(2):193–216. Available from: <https://doi.org/10.1257/jep.22.2.193>.
5. **Motorin V.I., Kenchadze D.D.** The Methods for Recoupling the Preliminary Quarterly Estimates of Product and Industry Outputs with Annual Production Data. *Voprosy Statistiki*. 2017;(10):3–12. (In Russ.)
6. European Commission (Eurostat). *Handbook on Quarterly National Accounts*. Luxembourg: Publications Office of the European Union; 2013. 312 p. Available from: <https://doi.org/10.2785/46080>.
7. European Commission (Eurostat). *Handbook on Quarterly National Accounts: Methods and Nomenclatures*. Luxembourg: Office for Official Publications of the European Communities; 1999. 436 p.
8. **Bloem A.M., Dippelsman R.J., Maehle N.O.** *Quarterly National Accounts Manual: Concepts, Data Sources, and Compilation*. Washington, D.C.: IMF; 2001. 210 p. (Russ. ed.: Blum E.M., Dippelsman R.Dzh., Mele N.E. Rukovodstvo po kvartal'nym natsional'nym schetam: kontseptsii, istochniki dannyykh i sostavlenie. Washington, DC: IMF; 2001. 222 p.)
9. **Dagum E.B., Cholette P.A.** *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*. Lecture Notes in Statistics # 186. New York: Springer Science+Business Media; 2006. 409 p.
10. International Monetary Fund. *Quarterly National Accounts Manual: 2017 edition*. Washington, D.C.: IMF; 2018. 272 p. (Russ. ed.: Rukovodstvo po kvartal'nym natsional'nym schetam: izdanie 2017 goda. Washington, D.C.: IMF; 2018. 315 p.)
11. **Miller R.E., Blair P.D.** *Input-Output Analysis: Foundations and Extensions: 2nd ed.* N.Y.: Cambridge University Press; 2009. 782 p. Available from: <https://doi.org/10.1017/CBO9780511626982>.
12. **Motorin V.** The Linear Matrix-Valued Cost Functions as a Source of Leontief and Ghosh Model. *Artha Vijnana: Journal of the Gokhale Institute of Politics and Economics*. 2017;59(4):291–301. Available from: <https://doi.org/10.21648/arthavij/2017/v59/i4/170795>.
13. **Motorin V.** Enhancing the Distance Minimization Methods of Matrix Updating Within a Homothetic Paradigm. *Journal of Economic Structures*. 2017;6(article 36):1–22. Available from: <https://doi.org/10.1186/s40008-017-0094-7>.

About the authors

Vladimir I. Motorin – Cand. Sci. (Econ.), Senior Research Fellow, Chief Expert of Economic Statistics Centre of Excellence, Department of Statistics and Data Analysis, National Research University Higher School of Economics (HSE University). 11, Pokrovsky Boulevard, Moscow, 109028, Russia. E-mail: motoriny@yandex.ru. ORCID: <https://orcid.org/0000-0003-0924-8275>.

Dmitry D. Kenchadze – Deputy Head, Federal State Statistics Service. 39, Miasnitskaya Str., Bldg. 1, 107450, Moscow, Russia. E-mail: Kenchadze@gks.ru.

Kirill A. Alekseev – Deputy Head of Department, National Accounts Department, Federal State Statistics Service. 39, Miasnitskaya Str., Bldg. 1, 107450, Moscow, Russia. E-mail: AlekseevKA@gks.ru.

Funding

The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE University).