
Financial University, Moscow, Russia; VByvshev@fa.ru; eLibrary SPIN: 6580-7089; ORCID: 0000-0002-8234-4936 Egor V. Pakhomov, Cand. Sc. (Economics), Deputy Director of the Department for Budget Planning, State Programmes and National Projects, Ministry of Economic Development of the Russian Federation, Moscow, Russia; PakhomovEV@economy.gov.ru; ORCID: 0009-0005-0848-5967

Abstract. An econometric assessment of the impact on Russia's gross domestic product (GDP) of the real primary deficit (budget impulse) and real budget expenditures is discussed. The main characteristic of the national economy is the annual growth rate of the country's real GDP in 2006–2021. An econometric model of this characteristic is constructed. The explanatory variables of the model include the values of the growth rate of the main factors of production and, in particular, the real primary budget deficit (budget impulse) and the growth rate of real budget expenditures. The main results of the work are as follows. 1. The growth rate of Russia's real GDP is significantly positively influenced by both the real primary budget surplus (budget impulse – as defined by the Ministry of Economic Development of Russia) and the growth rate of real budget expenditures. 2. Adverse events such as the global financial crisis, Western sanctions, and the pandemic have a significant negative impact on the Russian economy, reducing the growth rate of real GDP by an average of 2% per year.

Keywords: *econometric model of the growth rate of Russia's real GDP, real budget momentum, real budget expenditures, model of adaptation of the Russian economy to Western sanctions.*

Classification JEL: C50.

For reference: Byvshev V.A., Pakhomov E.V. (2024). Assessing the impact of budget momentum and budget expenditures on Russia's real GDP. *Economics of Contemporary Russia*, no. 3 (106). Pp. 24–37 (in Russian). DOI: 10.33293/1609-1442-2024-3(106)-24-37; EDN: LXGTVW

Manuscript received 05.03.2024

МЕТОД ПОСТРОЕНИЯ ВЕКТОРНЫХ АВТОРЕГРЕССИЙ ЛЮБОЙ СЛОЖНОСТИ

С.Г. Светульников¹

DOI: 10.33293/1609-1442-2024-3(106)-37-50

EDN: LXPIGW

Аннотация. Векторные авторегрессии являются одним из бурно развивающихся направлений многих областях современной науки. Они активно используются в моделировании и прогнозировании различных экономических процессов, чаще всего в моделировании фондового рынка и розничных цен. Их важнейшим преимуществом выступает возможность учета одновременного влияния моделируемых показателей не только от их прошлых значений, но и от прошлых значений других взаимосвязанных с ними показателей. Главная проблема, почему векторные авторегрессии не используются активно на практике (как они этого заслуживают), состоит в «проклятии размерности», которое заключается в квадратичном росте числа коэффициентов модели в зависимости от роста размерности моделируемого вектора. Это обстоятельство приводит к тому, что исследователи в разных областях современной науки вынуждены ограничивать размерность вектора, включая в модели только наиважнейшие либо снижая порядок авторегрессии. Попытки преодолеть «проклятие размерности» путем использования особых математических методов выливаются в существенное усложнение математического аппарата построения век-

© Светульников С.Г., 2024 г.

Светульников Сергей Геннадьевич, доктор экономических наук, профессор Высшей школы бизнес-инжиниринга Санкт-Петербургского политехнического университета Петра Великого, Санкт-Петербург, Россия; sergey@svetunkov.ru; eLibrary SPIN: 1772-7180

¹ Работа выполнена при финансовой поддержке РНФ, грант № 19-010-00610/19 «Теория, методы и методики прогнозирования экономического развития авторегрессионными моделями комплексных переменных».

торных авторегрессий, что не способствует расширению практики их применения. В статье предлагается использовать для этого поэтапный метод декомпозиции построения векторных авторегрессий любой размерности, который делает процесс построения этих моделей простым и доступным любому исследователю. Для проверки возможности применения этого метода на практике использовались ряды данных о динамике восьми основных отраслевых индексов Московской биржи. При этом было принято решение построить большую векторную авторегрессию порядка $p = 10$. С помощью метода наименьших квадратов всего было оценено 648 неизвестных коэффициентов этой модели. Верификация модели была подтверждена простыми авторегрессиями.

Ключевые слова: векторная авторегрессия, метод наименьших квадратов, размерность, лаг, метод поэтапной декомпозиции.

Классификация JEL: C22, C29, C53.

Для цитирования: Светульников С.Г. (2024). Метод построения векторных авторегрессий любой сложности // Экономическая наука современной России. № 3 (106). С. 37–50. DOI: 10.33293/1609-1442-2024-3(106)-37-50; EDN: LXPIGW

ВВЕДЕНИЕ

В последние годы на смену простым регрессиям приходят более сложные модели, одной из которых является моделирование и прогнозирование сложных стохастических процессов с помощью моделей векторных авторегрессий (VAR) или их модификаций (VARIMA) (Ord, Fildes, Kourentzes, 2017). Неоспоримым преимуществом векторных авторегрессий является возможность моделирования сложной многофакторной взаимозависимости выбранных показателей друг от друга с учетом возможной задержки их влияния во времени. Существенная часть стохастических процессов имеет именно такую природу, и в их числе процессы, протекающие на фондовых рынках.

Однако векторные авторегрессии так и не стали столь же распространенным ин-

струментом исследования, как простые одномерные авторегрессии. Вызвано это тем, что увеличение размера вектора k , который моделируется с помощью авторегрессий, приводит к нелинейному квадратичному росту числа коэффициентов авторегрессии, которые необходимо оценить по имеющимся статистическим данным. На определенном этапе увеличения числа моделируемых переменных k возникают вычислительные сложности, преодоление которых оказывается не под силу большинству исследователей. Например, если для построения векторной авторегрессии первого порядка размерности $k = 5$ необходимо оценить значения 25 неизвестных коэффициентов, то уже двукратный рост размерности вектора до $k = 10$ приводит к необходимости оценивать 100 неизвестных коэффициентов, а это даже при современном уровне развития вычислительной техники и при наличии многочисленных прикладных статистических программ является нетривиальной задачей для большинства исследователей.

Поэтому поиск методов упрощения процесса построения и оценивания векторных авторегрессий является важной научной и практической задачей.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Рассмотрим, как сегодня ученые строят векторные авторегрессии и оценивают их коэффициенты (Гельруд, Угрюмов, Рыбак, 2018; Зубарев, Кириллова, 2023). В общем виде $VAR_k(p)$ может быть представлена (см., например, (Kilian, Lütkepohl, 2017; Lütkepohl, 2005)) следующим образом:

$$\hat{Y}_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_p. \quad (1)$$

Здесь Y_t – k -мерный вектор переменных; A_0 – k -мерный вектор свободных коэффициентов, который часто приравнивается нулю; A_t –

$k \times k$ -мерные постоянные вещественные матрицы коэффициентов; p – лаг авторегрессии.

Очевидно, что фактический вектор моделируемых показателей Y_t описывается этой моделью $VAR(p)$ с некоторыми ошибками:

$$Y_t - \hat{Y}_t = E_t, \quad E_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \dots \\ \varepsilon_{kt} \end{pmatrix}. \quad (2)$$

Тогда, применяя к задаче оценивания коэффициентов модели (1) метод максимального правдоподобия и предполагая наличие нормального распределения ошибок, получим: для вычисления наилучших значений коэффициентов модели необходимо минимизировать значения дисперсионно-ковариационной матрицы (Геец и др., 2005):

$$\Lambda = \begin{pmatrix} \text{var}(\varepsilon_{1t}) & \text{cov}(\varepsilon_{1t}, \varepsilon_{2t}) & \dots & \text{cov}(\varepsilon_{1t}, \varepsilon_{kt}) \\ \text{cov}(\varepsilon_{2t}, \varepsilon_{1t}) & \text{var}(\varepsilon_{2t}) & & \text{cov}(\varepsilon_{2t}, \varepsilon_{kt}) \\ & & \ddots & \\ \text{cov}(\varepsilon_{kt}, \varepsilon_{1t}) & \text{cov}(\varepsilon_{kt}, \varepsilon_{2t}) & & \text{var}(\varepsilon_{kt}) \end{pmatrix}. \quad (3)$$

Здесь $\text{var}(\varepsilon_{it})$ – дисперсия ошибок аппроксимации модели i ; $\text{cov}(\varepsilon_{it}, \varepsilon_{jt})$ – ковариация ошибок аппроксимации моделей i и j .

Переменными, влияющими на дисперсионно-ковариационную матрицу (3), выступают коэффициенты $VAR(p)$, число которых будет равно k^2p . При большом размере k моделируемого вектора число этих коэффициентов становится столь значительным, что решить задачу подбора лучших значений коэффициентов для минимизации (3) становится сложно.

Для понимания сложности решаемой задачи размер вектора авторегрессии является ключевым показателем, поэтому следует указывать его в обозначении модели. Тогда запись $VAR^k(p)$ следует понимать как векторную авторегрессию порядка p для вектора размерностью k (Светульников, Баженова, Рыбак, 2022). Это обозначение будем использовать в дальнейшем.

Почему размер моделируемого вектора решающим образом влияет на уровень

сложности задачи? Это можно понять, если, например, в модели $VAR^5(5)$ увеличить в два раза размер моделируемого вектора и использовать $VAR^{10}(5)$. Для применения на практике исходной модели исследователю необходимо было оценить $k^2p = 125$ коэффициентов, что является непростой, но решаемой задачей. Но удвоение вектора до $k = 10$ переменных приводит к необходимости оценить значения $k^2p = 500$ неизвестных коэффициентов. А вот решение подобной задачи является нетривиальным, и ее решение под силу только высококвалифицированным специалистам.

Именно вычислительные сложности приводят к тому, что применяемые на практике модели $VAR^k(p)$ в основном используются для векторов невысокого порядка ($k < 6$) и с малым числом лагов ($p < 4$). Так, Ж.А. Алиаскарова, А.Б. Асадулаев и В.Ю. Пашкус построили модель $VAR^2(4)$ (Алиаскарова, Асадулаев, Пашкус, 2020), т.е. вычислили 16 неизвестных коэффициентов. Н. Маматова использовала $VAR^3(1)$ (Маматова, 2015), вычисляя значения 9 неизвестных коэффициентов. Д.Н. Малов использовал несколько моделей $VAR_k(1)$ при разных размерностях векторов от $k = 3$ до $k = 5$, т.е. оценивая от 9 до 25 неизвестных коэффициентов (Малов, 2019). К. Погосян, моделируя экономику Армении, использовал $VAR^3(4)$ (Погосян, 2015) с 36 коэффициентами.

Большая часть зарубежных исследователей также использует $VAR^k(p)$ небольшой размерности. В своей докторской диссертации А. Barrett использовал при исследовании криптовалют $VAR^2(2)$ (Barrett, 2021), т.е. оценивал 8 неизвестных коэффициентов. Исследуя выбросы CO_2 в Китае, Бин Ху и Бокьянг Дин построили модель $VAR^3(4)$, которая содержит 36 неизвестных коэффициентов (Ху, Boqiang, 2016). Н.Н. Каролин и М. Шеррис, изучая смертность населения, использовали векторные авторегрессии $VAR^5(2)$ (Carolyn, Sherris, 2020), оценивая 50 коэффициентов. Краткосрочное прогнозирование дорожного движения на автомагистралях США С.Р. Чандра и Х. Аль-Дик осуществили с помощью

$VAR^5(1)$ (Chandra, Al-Deek, 2009) 25 неизвестных коэффициентов. Моделирование климатических данных о количестве осадков, влажности, максимальной и минимальной температурах и продолжительности солнечного излучения в день с использованием подхода векторной авторегрессии интегрированного скользящего среднего $VARIMA^5(1, 1, 1)$ осуществили Ч. Фатрианти, С.М. Бельвавин, М. Рийана и Р. Амин (Fitrianti et al., 2019). При этом они оценивали 25 коэффициентов векторной авторегрессии и 25 коэффициентов векторного скользящего среднего – всего 50 коэффициентов. К. Гарсиа-Мартос с соавторами с помощью $VAR^3(1)$ промоделировал и спрогнозировал мировые цены на ископаемое топливо, цены на квоты на выбросы CO_2 , а также цены на электроэнергию (Garcia-Martos, Rodriguez, Sanchez, 2013). Как видно, эти ученые оценивали 9 неизвестных коэффициентов. Данные по инфляции и процентным ставкам в Индонезии М.Ч. Джасмавати и В. Пенерапан моделировали с помощью довольно редкой модели $VARIMA^4(0,2,2)$, где авторегрессия данных не учитывается, а используется только векторная скользящая средняя ошибок второго порядка (Jusmawati, Penegarpan, 2020). Можно заметить, что всего ими было оценено 32 неизвестных коэффициента. Индонезийские ученые А. Меймела с соавторами построили $VARIMA^3(1,1,1)$ для моделирования и прогнозирования смертности от COVID-19 (Meimela et al., 2021), т.е. они оценили 18 неизвестных коэффициентов на статистических данных. Метеорологическое прогнозирование четырех показателей осадков с использованием $VARIMA^4(1,1,2)$ для использования этих прогнозов рыбаками осуществили А. Русийана с соавторами (Rusyana et al., 2020). При этом они оценивали 48 неизвестных коэффициентов. В долгосрочном оптимальном управлении гидроветряными гибридными системами Китая Цанг И с соавторами предложил использовать двумерную векторную авторегрессию прогнозирования стока воды и напора ветра (Zhang et al., 2021). Число коэффициентов таких моделей не пре-

вышает двух десятков при лаге авторегрессии, равном пяти.

Модель $VARIMA^3(1,1,1)$ была построена Ф.А. Цейлер с соавторами для изучения взаимосвязи между медленноволновыми колебаниями внутричерепного давления, среднего артериального давления и $PbtO_2$ 47 пациентов (Zeiler, Ercole, Cabeleira, 2020). Они смогли оценить, как можно заметить, 18 неизвестных коэффициентов. Модель VARIMA была построена китайскими учеными для долгосрочного управления гидро-ветровой гибридной системой на юго-западе Китая (Yi et al., 2021). При этом оценивалось 27 неизвестных коэффициентов.

В реальных ситуациях, подобным тем, которые изложены в этих публикациях, число моделируемых переменных должно было быть больше, но это существенно увеличивает размерность задачи, и поэтому число переменных ограничивается. Поэтому ученые непрерывно искали пути упрощения задачи, чтобы за счет подобных процедур упрощения преодолеть «проклятие размерности». И отдельные успешные решения были получены. Так, например, одним из методов снижения размерности задачи может быть использование метода главных компонент (Barrett, 2021). В два раза снизить число оцениваемых коэффициентов возможно с помощью представления вектора в комплекснозначной форме (Светульников, Баженова, Лукаш, 2022). Можно использовать различные процедуры исключения из модели тех ее составляющих, влияние которых на результат минимально (Carolyn, Sherris, 2020). Появились работы, которые синтезируют с векторными авторегрессиями нейронные сети (Lusia, Ambarwati, 2018; Schmidhuber, 2015) и др. Но эти предложения, снижая размерность задачи, повышают ее сложность за счет привлечения непростого математического аппарата, поэтому они не нашли широкого применения на практике.

Имеется другая возможность упростить задачу, если оценивать не одновременно коэффициенты всех уравнений, входящих в модель (1), а только нескольких из них. Действительно при

минимизации детерминанта дисперсионно-ковариационной матрицы (3) учитываются не только дисперсии ошибок аппроксимации моделью каждой составляющей y_{it} моделируемого вектора, но и ковариации ошибок аппроксимации $\text{cov}(\varepsilon_{it}, \varepsilon_{jt})$. Из теории вероятностей известно, что ковариация двух независимых случайных переменных равна нулю. Если модель построена верно и она правильно описывает сложный процесс, то ошибки аппроксимации не должны быть коррелированы друг с другом, т.е. ковариация между ними должна быть равна нулю. Очевидно, что это будет выполняться на практике не всегда, но предположение о том, что все ковариации дисперсионно-ковариационной матрицы (3) могут быть приравнены нулю, представляется вполне рациональным. В таком случае дисперсионно-ковариационная матрица (3) примет вид

$$\Lambda = \begin{pmatrix} \text{var}(\varepsilon_{1t}) & 0 & \dots & 0 \\ 0 & \text{var}(\varepsilon_{2t}) & & 0 \\ & & \ddots & \\ 0 & 0 & & \text{var}(\varepsilon_{kt}) \end{pmatrix}. \quad (4)$$

Здесь $\text{var}(\varepsilon_{it})$ – дисперсия каждого отдельного моделируемого показателя.

И минимизация детерминанта такой матрицы даст точно такие же результаты оценок коэффициентов модели, как и результаты минимизации всех отдельных дисперсий, составляющих эту матрицу. Но в последнем случае переменными задачи оптимизации выступают не все k^2p коэффициенты одновременно, а только kp коэффициентов каждой строки модели (1) для переменной i в y_{it} , поскольку минимизируется дисперсия ошибки ее аппроксимации:

$$\text{var}(\varepsilon_{it}) = \frac{1}{T} \sum_t (y_{it} - \hat{y}_{it})^2, \quad (5)$$

где расчетное значение представляет собой простую авторегрессию

$$\hat{y}_{it} = a_{0i} + a_{1i}y_{1(t-1)} + a_{2i}y_{2(t-1)} + \dots + a_{pi}y_{p(t-p)}. \quad (6)$$

Поскольку число таких уравнений равно k , то эту задачу надо будет решать k раз, и в итоге с помощью МНК будет оценено k^2p неизвестных коэффициентов.

Можно убедиться, что решается та же самая исходная задача, но решается она значительно проще, поскольку каждый раз при поиске оптимальных значений неизвестных коэффициентов $a_{ji}, j = 1, 2, \dots, p, i = 1, 2, \dots, k$ необходимо оценить только kp коэффициентов. Эту форму векторной авторегрессии иногда называют «приведенной формой» (Дорохов, 2008).

Переформулирование задачи минимизации (3) в задачу (4) существенно расширило возможности использования в исследованиях векторных авторегрессий больших размерностей. И в результате этого появились научные публикации, в которых указывалось на успешное применение более сложных, чем ранее, моделей.

Так, например, Е.В. Дорохов построил $VAR^k(p)$ с параметрами: $k = 29, p = 1$ (Дорохов, 2008). С помощью МНК он 30 раз оценивал 30 неизвестных коэффициентов (в его модели присутствует вектор свободных членов), что является вполне посильной задачей. О.Н. Салманов, В.М. Заернюк и О.А. Лопатина провели исследование, используя $VAR^7(2)$ (Салманов, Заернюк, Лопатина, 2016), т.е. они семь раз последовательно оценивали по 14 неизвестных коэффициентов. Е.И. Суханова и С.Ю. Ширнаева построили $VAR^6(3)$ (Суханова, Ширнаева, 2014), последовательно оценивая по 18 неизвестных коэффициентов, а Д.В. Шимановский и Е.А. Третьякова построили $VAR^6(2)$ (Шимановский, Третьякова, 2020), т.е. они шесть раз оценивали одновременно по 12 неизвестных коэффициентов модели.

Новые индексы цереброваскулярной реактивности на основе $VARIMA^3(4,1,4)$, где оценивалось 72 неизвестных коэффициента, предложила использовать группа исследователей (Thelin et al., 2020). Другие ученые (Ahelegbey, Billio, Casarin, 2016) использовали $VAR^6(4)$, оценивая 144 неизвестных коэффициента. При исследовании взаимосвязи

фондовых индексов стран – членов БРИКС, индексов других бирж и некоторых криптовалют была построена $VAR^8(1)$. При этом оценивалось 64 неизвестных коэффициента модели. Изучение 11 взаимосвязанных показателей выброса вредных веществ в атмосферу с помощью $VAR^{11}(1)$ осуществили Д. Олсон, Т. Ридел и другие ученые (Olson et al., 2021) – авторы оценивали 121 неизвестный коэффициент модели.

Отмечая заметное снижение сложности решения поставленной задачи за счет приравнивания нулю ковариаций ошибок аппроксимации, следует все же отметить, что это частичное решение задачи. Действительно, используя представление (4) для оценивания параметров модели $VAR^{10}(10)$, содержащей $k^2p = 1000$ коэффициентов, получим, что одновременно необходимо оценить не 1000 неизвестных коэффициентов, а 100. Это, конечно же, проще, но все равно затруднительно, поскольку не каждый исследователь может решить задачу одновременной оценки 100 неизвестных коэффициентов. Именно поэтому исследователи получили возможность увеличивать число моделируемых показателей в векторной авторегрессии, но все-таки не до бесконечности.

Еще в 1970-е годы директор Института кибернетики АН УССР академик А.Г. Ивахненко успешно работал над проблемой моделирования сложных процессов, в том числе с помощью рядов Винера (полиномов Колмогорова–Габора) (Ивахненко, Мюллер, 1985):

$$y = a_0 + \sum_{i=1}^m a_i x_i + \sum_{i=1}^m \sum_{j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m a_{ijk} x_i x_j x_k + \dots, \quad (7)$$

где y – результирующая переменная; а x_i – переменные, влияющие на результат.

Этот полином хорошо описывает нелинейную связь между переменной y и факторами x_i . У этой задачи есть своя сложность, когда в полином включаются два фактора, число коэффициентов полинома оказывается

равным шести. Если же число объясняющих переменных равно трем, то число членов полинома возрастает уже до 20 и т.д.

Получается, что количество членов полинома Колмогорова–Габора (и коэффициентов при них) нелинейно возрастает и при числе факторов m и показателе степени q составит (Мехович, Ахиезер, Дунаевская, 2014; Петров и др., 2020)

$$k = C_{m+q}^m = \frac{(m+q)!}{m!q!}. \quad (8)$$

В этом случае также можно наблюдать «проклятие размерности».

А.Г. Ивахненко предложил свой метод снижения влияния «проклятия размерности» на результаты, назвав его методом группового учета аргументов (МГУА). Этот метод построения полинома Колмогорова–Габора используется в настоящее время для решения некоторых частных задач (Баласанян, Геворгян, 2016; Белов, Чистякова, 2008; Дьячков, 2017). Мы не будем подробно разбирать этот метод, поскольку в нем есть не имеющие к нашей задаче аспекты, а укажем лишь на те его существенные черты, которые помогут нам при оценке коэффициентов $VAR_k(p)$.

А.Г. Ивахненко продемонстрировал суть МГУА на примере оценки методом наименьших квадратов коэффициентов полинома Колмогорова–Габора для трех объясняющих факторов:

$$\begin{aligned} y = & a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_1^2 + \\ & + a_5 x_2^2 + a_6 x_3^2 + a_7 x_1 x_2 + a_8 x_1 x_3 + \\ & + a_9 x_2 x_3 + a_{10} x_1^3 + a_{11} x_2^3 + a_{12} x_3^3 + \\ & + a_{13} x_1^2 x_2 + a_{14} x_1^2 x_3 + a_{15} x_1 x_2^2 + \\ & + a_{16} x_2^2 x_3 + a_{17} x_1 x_3^2 + a_{18} x_2 x_3^2 + \\ & + a_{19} x_1 x_2 x_3. \end{aligned} \quad (9)$$

Задачу оценки этих 20 неизвестных коэффициентов он разбил на два этапа. На первом этапе с помощью МНК оцениваются коэффициенты трех более простых частных полиномов, а именно

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1^2 + b_4x_2^2 + b_6x_1x_2 + \varepsilon_1 = \hat{y}_1 + \varepsilon_1; \quad (10)$$

$$y = c_0 + c_1x_1 + c_2x_3 + c_3x_1^2 + c_4x_3^2 + c_6x_1x_3 + \varepsilon_2 = \hat{y}_2 + \varepsilon_2; \quad (11)$$

$$y = d_0 + d_1x_2 + d_2x_3 + d_3x_2^2 + d_4x_3^2 + d_6x_2x_3 + \varepsilon_3 = \hat{y}_3 + \varepsilon_3. \quad (12)$$

Конечно же, ошибки аппроксимации ε_1 , ε_2 и ε_3 имеют минимально возможную дисперсию.

На втором этапе было предложено использовать эти расчетные значения частных полиномов и подставлять их в уравнение

$$y = e_0 + e_1\hat{y}_1 + e_2\hat{y}_2 + e_3\hat{y}_3 + e_4\hat{y}_1\hat{y}_2\hat{y}_3 + \varepsilon, \quad (13)$$

коэффициенты которого также предлагалось оценивать с помощью МНК.

Легко заметить, что в случае, когда вместо расчетных значений показателей $\hat{y}_1, \hat{y}_2, \hat{y}_3$ в (13) подставить формулы частных полиномов (10)–(12), раскрыть скобки и сгруппировать, то будет получен полный полином Колмогорова–Габора (9), в котором, правда, появляются члены полинома, не соответствующие членам полинома Колмогорова–Габора (9), например члены с переменными в четвертой степени x_1^4, x_2^4 и x_3^4 .

Мы не будем использовать полином Колмогорова–Габора, а воспользуемся самой идеей разбиения общей задачи на ряд частных подзадач для построения векторных авторегрессий. Сам А.Г. Ивахненко называл этот подход по-разному: многорядным, многоэтапным, многоуровневым, и он являлся важной составляющей метода группового учета аргументов (МГУА), который, собственно говоря, А.Г. Ивахненко и считал основным. Нас МГУА не интересует, поскольку он применялся для решения других задач, а вот идею декомпозиции задачи оценивания параметров модели возьмем за основу и расширим ее до уровня метода поэтапной декомпозиции SDM (stepwise decomposition method).

С помощью SDM будем находить значения коэффициентов модели $VAR_k(p)$, минимизируя (4), т.е. последовательно находя коэффициенты k моделей (6).

Запишем (6) более подробно, сгруппировав ее составляющие в отдельные группы, отличающиеся друг от друга по величине лага (возможны и другие варианты группировки):

$$\begin{aligned} y_{it} = & a_{i0} + a_{i11}y_{1(t-1)} + a_{i21}y_{2(t-1)} + \\ & + \dots + a_{ik1}y_{k(t-1)} + a_{i12}y_{1(t-2)} + a_{i22}y_{2(t-2)} + \\ & + \dots + a_{ik2}y_{k(t-2)} + \dots + a_{i1p}y_{1(t-p)} + \\ & + a_{i2p}y_{2(t-p)} + \dots + a_{ikp}y_{k(t-p)}. \end{aligned} \quad (14)$$

Как видно, здесь необходимо оценить kp неизвестных коэффициентов, где k – число переменных, включенных в моделируемый вектор; p – величина лага.

Оценивание этих коэффициентов может быть осуществлено в два этапа. Первый этап – строятся многофакторные модели зависимости данного показателя от всех других показателей, но сгруппированных по признаку одинаковости величины лага: сначала для лага, равного единице, затем для лага, равного двум, и т.д. до лага, равного p :

$$\begin{aligned} \hat{y}_{i(1)} = & b_0 + b_1y_{1(t-1)} + b_2y_{2(t-1)} + \dots + b_ky_{k(t-1)}, \\ \hat{y}_{i(2)} = & c_0 + c_1y_{1(t-2)} + c_2y_{2(t-2)} + \dots + c_ky_{k(t-2)}, \\ & \dots \\ \hat{y}_{i(p)} = & z_0 + z_1y_{1(t-p)} + z_2y_{2(t-p)} + \dots + z_ky_{k(t-p)}, \end{aligned} \quad (15)$$

коэффициенты этих моделей оцениваются с помощью МНК, предполагая наличие нормального распределения остатков.

Число коэффициентов каждой из моделей (15) будет соответствовать размеру вектора k плюс 1 (поскольку есть свободный член), а число таких уравнений – величине лага p . Если, например, необходимо построить $VAR^{10}(10)$, то число неизвестных коэффициентов в каждом уравнении будет равно $k + 1 = 11$, а всего необходимо построить $p = 10$ таких уравнений. А эту задачу решать уже легко.

Теперь переходим ко второму этапу оценки коэффициентов уравнения (14). Для

этого расчетные значения переменной y_{it} , вычисленные по каждой из моделей (15), следует подставить в уравнение, состоящее из p членов:

$$\hat{y}_{it} = \alpha_0 + \alpha_1 \hat{y}_{i(1)} + \alpha_2 \hat{y}_{i(2)} + \dots + \alpha_p \hat{y}_{i(p)}. \quad (16)$$

После этого оцениваются коэффициенты этой модели второго этапа.

Если теперь подставить в (16) модели (15), раскрыть скобки и сгруппировать члены, то можно найти значения коэффициентов исходной модели (14) через коэффициенты моделей (15) и (16). Например, значение свободного члена ai_0 вычисляется так:

$$ai_0 = \alpha_0 + \alpha_1 b_0 + \alpha_2 c_0 + \dots + \alpha_p z_0. \quad (17)$$

Значения всех коэффициентов пропорциональности вычисляются проще. Так, коэффициент a_{i11} рассчитывается так:

$$a_{i11} = \alpha_1 b_1. \quad (18)$$

Коэффициент a_{i21} модели (14) находится так:

$$a_{i21} = \alpha_1 b_2. \quad (19)$$

И для любого коэффициента a_{ijl} вычисления сводятся к простому перемножению:

$$a_{ijl} = \alpha_l b_j. \quad (20)$$

Как видно, метод поэтапной декомпозиции SDM представляет собой линейную суперпозицию линейных функций, а это значит, что оценки МНК, полученные непосредственно к исходной модели (14), и оценки МНК, полученные методом SDM, совпадут.

Рассмотрим теперь возможность применения SDM для построения сверхбольших векторных авторегрессий. Число этапов применения метода SDM может быть и больше, чем два, и определяется размерностью задачи.

Пусть, например, исследователь предполагает построить сверхбольшую векторную авторегрессию $VAR^{100}(200)$ без свободного члена (его значения равны нулю). Такая мо-

дель будет содержать 200 векторов и 200 квадратных матриц коэффициентов при них, а сами квадратные матрицы коэффициентов будут состоять из 10 тыс. неизвестных коэффициентов. Всего предстоит оценить 2 млн неизвестных коэффициентов. Для построения такой большой модели вновь следует считать ковариации матрицы (3) равными нулю и находить коэффициенты не одновременно, а по отдельным строкам, моделирующим показатели y_{it} , $i = 1, 2, 3, \dots, 100$. Но даже в этом случае оценивать необходимо модели, содержащие $100 \times 200 = 20\,000$ неизвестных коэффициентов, что при традиционном подходе является неразрешимой задачей, но для SDM эта задача вполне по силам.

Пусть для определенности исследователю удобно одновременно оценивать по 20 коэффициентов, не больше.

Тогда на первом этапе модель y_{it} , содержащая 20 тыс. коэффициентов, разбивается на 1000 моделей, содержащих по 20 коэффициентов, – так, как это было показано в (14), (15). Найти значения коэффициентов этих моделей с помощью МНК является заурядной задачей.

Далее, на втором этапе, эти 1000 моделей вновь группируются в 50 групп по 20 моделей в каждой. Для каждой из этих 50 групп строится модель по принципу (16) с использованием расчетных значений 20 моделей, входящих в группу.

На третьем этапе 50 получившихся моделей вновь разбиваются, но уже на три группы – две по 20 и одна по 10 моделей, поскольку мы предположили, что исследователь может одновременно оценить не более 20 коэффициентов модели. И вновь для каждой из трех групп третьего этапа вычисляется единая модель по принципу (16).

Расчет коэффициентов модели завершен.

Если у исследователя будет желание на основе моделей этих четырех этапов построить общую модель $VAR^{100}(200)$ в явном виде с 20 тыс. коэффициентов, то он может это сделать, воспользовавшись правилами (15)–(18).

Но для практических нужд этого делать не обязательно. Следует сохранить коэффици-

енты, полученные на каждом этапе, и использовать их для исследований, например, получив новые значения $y_{i(t+\tau)}$. Вначале их следует подставить в модели первого этапа и вычислять расчетные значения. Далее расчетные значения первого этапа следует подставлять в модели второго этапа и вновь вычислять промежуточные расчетные значения модели. Их следует подставить в модели, полученные на третьем этапе, а результаты этих расчетов можно подставить в последнюю модель четвертого этапа и получить итоговое значение моделируемого показателя $y_{i(t+\tau)}$.

Как видно, предлагаемая многоэтапная процедура SDM построения $VAR^k(p)$ не требует привлечения каких-либо особых сложных методов, и она под силу экономисту, владеющему математической статистикой на среднем уровне.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Метод поэтапной декомпозиции SDM следует апробировать на каком-либо примере, прежде чем рекомендовать его для широкого применения на практике. Таким объектом моделирования был выбран рынок Московской биржи, на которой были выбраны данные об изменении основных отраслевых индексов этой биржи: нефти и газа y_{1t} , финансов y_{2t} , металла и добычи полезных ископаемых y_{3t} , электроэнергетики y_{4t} , потребительского сектора y_{5t} , химии и нефтехимии y_{6t} , транспорта y_{7t} и телекоммуникаций y_{8t} . Были выбраны еженедельные данные по этим индексам начиная с 01.05.2016 по 31.10.2021 (всего – 289 наблюдений).

Для того чтобы избежать проблем в размерности, при построении подобных моделей прибегают к нормализации исходных данных, что и было выполнено.

Поскольку была поставлена задача проверить возможность применения SDM для построения векторных авторегрессий большого размера, а не выбора лучшей модели для

моделирования этих индексов, было принято решение построить модель $VAR^k(p)$ вектора размерности $k = 8$ и с лагом, равным $p = 10$:

$$\hat{Y}_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + A_3 Y_{t-3} + \dots + A_{10} Y_{t-10}.$$

Здесь предстояло вычислить уже $k^2 p + k = 8^2 \times 10 + 8 = 648$ неизвестных коэффициентов.

Модель является большой, но не сверхбольшой, поэтому SDM применялась в два этапа.

На первом этапе строилось 10 многофакторных моделей типа (15), сформированных по величине лага. Так, например, индекс нефтегазового сектора в момент времени t вначале определялся как зависимость от цен на все индексы в предшествующий момент времени $(t - 1)$:

$$\hat{y}_{1t(1)} = b_0 + b_1 y_{2(t-1)} + b_2 y_{3(t-1)} + b_3 y_{4(t-1)} + b_4 y_{4(t-1)} + b_5 y_{5(t-1)} + b_6 y_{6(t-1)} + b_7 y_{7(t-1)} + b_8 y_{8(t-1)}. \quad (21)$$

Это элементарная задача, и девять коэффициентов модели находились без труда:

$$\hat{y}_{1t(1)} = 0,011 + 0,981 y_{2(t-1)} - 0,078 y_{3(t-1)} + 0,063 y_{4(t-1)} - 0,012 y_{4(t-1)} + 0,037 y_{5(t-1)} + 0,055 y_{6(t-1)} - 0,004 y_{7(t-1)} - 0,063 y_{8(t-1)}.$$

Затем оценивались коэффициенты зависимости этого индекса y_{1t} от восьми индексов, но с лагом, равным двум $(t - 2)$:

$$\hat{y}_{1t(2)} = c_0 + c_1 y_{2(t-2)} + c_2 y_{3(t-2)} + c_3 y_{4(t-2)} + c_4 y_{4(t-2)} + c_5 y_{5(t-2)} + c_6 y_{6(t-2)} + c_7 y_{7(t-2)} + c_8 y_{8(t-2)}. \quad (22)$$

И так далее – до модели зависимости y_{1t} от индексов, сдвинутых на максимально заданный лаг $p = 10$:

$$\hat{y}_{1t(10)} = z_0 + z_1 y_{2(t-10)} + z_2 y_{3(t-10)} + z_3 y_{4(t-10)} + z_4 y_{4(t-10)} + z_5 y_{5(t-10)} + z_6 y_{6(t-10)} + z_7 y_{7(t-10)} + z_8 y_{8(t-10)}. \quad (23)$$

Как видно, первый этап построения модели с помощью SDM не вызывает никаких затруднений – коэффициенты каждой модели легко оцениваются с помощью МНК.

На втором этапе SDM расчетные значения 10 моделей (21)–(22) подставлялись в общую модель (16), которая для нашего случая имеет вид

$$\hat{y}_{1t} = \alpha_0 + \alpha_1 \hat{y}_{1t(1)} + \alpha_2 \hat{y}_{1t(2)} + \alpha_3 \hat{y}_{1t(3)} + \dots + \alpha_9 \hat{y}_{1t(9)} + \alpha_{10} \hat{y}_{1t(10)}. \quad (24)$$

Оценка этих 11 коэффициентов по имеющимся фактическим данным и расчетным значениям также представляет собой заурядную задачу:

$$\begin{aligned} \hat{y}_{1t} = & 0,001 + 1,094y_{1t(1)} - 0,177y_{1t(2)} - \\ & - 0,044y_{1t(3)} + 0,187y_{1t(4)} + 0,065y_{1t(5)} - \\ & - 0,255y_{1t(6)} + 0,230y_{1t(7)} - 0,152y_{1t(8)} + \\ & + 0,047y_{1t(9)} + 0,004y_{1t(10)}. \end{aligned}$$

Теперь по рассчитанным значениям коэффициентов (24) и коэффициентам моделей первого этапа (21)–(23) определялись те коэффициенты векторной авторегрессии, которые описывают поведение стоимости индекса нефтегазового сектора в зависимости от изменения всех восьми индексов в предшествующие моменты времени – от предыдущей недели до сдвинутых на 10 недель назад. Всего, как можно убедиться, был оценен 81 коэффициент.

Аналогично оценивался 81 коэффициент зависимости второго индекса y_{2t} от других индексов, 81 коэффициент зависимости третьего индекса y_{3t} от восьми других индексов и т.д. – всего восемь уравнений, которые, будучи представлены в матричной форме, соответствуют общей модели (1).

Как видно, процесс построения $VAR^8(10)$ оказался довольно простым, но рутинным. Его формализация на наиболее распространенных сегодня среди ученых языках программирования R или Python делает задачу построения модели авторегрессии $VAR^k(p)$ любой степени сложности доступной для многочисленных исследователей.

Для того чтобы убедиться в правильности построенной модели, нужно ее верифицировать. Для этой цели были использованы модели простых авторегрессий $AR(p)$ десятого порядка:

$$\hat{y}_{it} = a_{i0} + \sum_{p=1}^{10} a_{ip} y_{i(t-p)}. \quad (25)$$

Всего было построено восемь моделей авторегрессий $AR_i(10)$, $i = 1, 2, \dots, 8$, – по числу имеющихся переменных $k = 8$. После этого вычислялась ошибка аппроксимации по всем данным для каждого из 10 индексов по общей модели $VAR^8(10)$ и по соответствующей этим индексам моделям авторегрессий $AR_i(10)$. Обобщенной оценкой точности описания каждого ряда индексов служила средняя абсолютная ошибка аппроксимации $MAPE$:

$$MAPE = \frac{1}{T} \sum_{t=1}^T |y_{it} - \hat{y}_{it}|. \quad (26)$$

Значения ошибок аппроксимации по модели векторной авторегрессии и по восьми моделям простых авторегрессий сведены в таблицу.

Как следует из данных таблицы, SDM подтверждает свою значимость. С его помощью была построена модель векторной авторегрессии, содержащая 648 коэффициентов. Усложнение модели практически всегда приводит к повышению точности аппроксимации. Вот и на этот раз все индексы немного точнее описываются с помощью $VAR^k(p)$, чем это делает простая авторегрессия $AR(p)$, за исключением индекса транспортного сектора, который одинаково точно описывается как простой авторегрессией, так и векторной. Выбор лучшей модели векторной авторегрессии теперь сделать просто – последовательно увеличивая сложность модели и оценивая статистические характеристики модели, с помощью одного из известных критериев выбора авторегрессии выбирается модель оптимальной сложности.

Таблица
Результаты верификации векторной авторегрессии

Статистические характеристики	Нефть и газ y_{1t}	Финансы y_{2t}	Металл и добыча y_{3t}	Электроэнергетика y_{4t}	Потребительский сектор y_{5t}	Химия и нефтехимия y_{6t}	Транспорт y_{7t}	Телекоммуникации y_{8t}
Значение $MAPE$ для $AR_L(10)$	0,098	0,081	0,063	0,151	0,085	0,072	0,119	0,138
Значение $MAPE$ для $VAR^8(10)$	0,095	0,078	0,062	0,148	0,082	0,070	0,119	0,135
Процент повышения точности аппроксимации $VAR^8(10)$ по сравнению с $AR_L(10)$	2,879	4,664	1,846	1,512	3,441	2,677	0,000	1,595

ЗАКЛЮЧЕНИЕ

Изложенный в статье метод построения векторных авторегрессий любых размерностей решает пресловутую «проблему размерности» путем поэтапной оценки коэффициентов этих моделей. При использовании SDM исчезают многочисленные вычислительные трудности, характерные для задачи при использовании векторов большой размерности и большом числе лагов. Открывающиеся перспективы, конечно же, позволят исследователям широко использовать векторные авторегрессии в практике моделирования различных задач. Среди таких перспектив – построение и больших нелинейных моделей векторных авторегрессий $NVAR_k(p)$, и больших векторных авторегрессий – скользящего среднего $VARIMa_k(p, d, q)$, и других модификаций векторных авторегрессий.

Список литературы / References

- Алиаскарова Ж.А., Асадулаев А.Б., Пашкус В.Ю. (2020). Прогнозирование динамики инвестиций в основной капитал и валовую добавленную стоимость на основе моделей VAR и VECM // Проблемы современной экономики. № 4 (76). С. 41–45. [Aliaskarova Zh.A., Asadulaev A.B., Pashkus V.Yu. (2020). Forecasting the dynamics of investments in fixed capital and gross added value based on VAR and VECM models. *Problems of Modern Economics*, no. 4 (76), pp. 41–45 (in Russian).]
- Баласанян С.Ш., Геворгян Э.М. (2016). Сравнительный анализ методов регрессии и группового учета аргументов при моделировании процессов переработки полезных ископаемых // Известия Томского политехнического университета. Инжиниринг георесурсов. Т. 327. № 4. С. 23–34. [Balasanyan S.Sh., Gevorgyan E.M. (2016). Comparative analysis of regression methods and group accounting of arguments in modeling mineral processing processes. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, vol. 327, no. 4, pp. 23–34 (in Russian).]
- Белов В.В., Чистякова В.И. (2008). Моделирование и прогнозирование бизнес-процессов с помощью алгоритмов самоорганизации формальных описаний // Бизнес-информатика. № 4 (06). С. 37–45. [Belov V.V., Chistyakova V.I. (2008). Modeling and forecasting business processes using algorithms for self-organization of formal descriptions. *Business Informatics*, no. 4 (06), pp. 37–45 (in Russian).]
- Геець В.М., Клебанова Т.С. и др. (2005). Модели и методы социально-экономического прогнозирования: учебник. Харьков: ВД «ИНЖЕК». 396 с. [Geets V.M., Klebanova T.S. et al. (2005). *Models and methods of socio-economic forecasting: textbook*. Kharkov: VD “INZHEK”. 396 p. (in Ukrainian).]
- Гельруд Я.Д., Угрюмов Е.А., Рыбак В.Л. (2018). Векторная модель авторегрессии показателей производственной деятельности строительного предприятия // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. Т. 7. № 3. С. 19–30. [Gelrud Ya.D.,

- Ugryumov E.A., Rybak V.L. (2018). Vector model of autoregression of production activity indicators of a construction enterprise. *Bulletin of SUSU. Series: Computational Mathematics and Computer Science*, vol. 7, no. 3, pp. 19–30 (in Russian).]
- Дорохов Е.В. (2008). Статистический подход к изучению прогнозирования индекса РТС на основе методов векторной авторегрессии и коинтеграции // *Финансы и бизнес*. № 1. С. 85–110. [Dorokhov E.V. (2008). Statistical approach to the study of forecasting the RTS index based on vector autoregression and cointegration methods. *Finance and Business*, no. 1, pp. 85–110 (in Russian).]
- Дьячков М.Ю. (2017). Индуктивное моделирование объектов и явлений методом группового учета аргументов: недостатки и способы их устранения // *Вестник Российского университета дружбы народов. Серия: Математика. Информатика. Физика*. Т. 25. № 4. С. 323–330. [Dyachkov M.Yu. (2017). Inductive modeling of objects and phenomena by the method of group accounting of arguments: shortcomings and ways to eliminate them. *RUDN Journal. Series: Mathematics. Computer science. Physics*, vol. 25, no. 4, pp. 323–330 (in Russian).]
- Зубарев А.В., Кириллова М.А. (2023). Построение модели GVAR для российской экономики // *Экономический журнал Высшей школы экономики*. № 1. С. 9–32. [Zubarev A.V., Kirillova M.A. (2023). Construction of the GVAR model for the Russian economy. *HSE Economic Journal*, no. 1, pp. 9–32 (in Russian).]
- Ивахненко А.Г., Мюллер Й.А. (1985). Самоорганизация прогнозирующих моделей. Киев: Техника; Берлин: ФЕБ Ферлаг Техник, 1985. 223 с. [Ivakhnenko A.G., Muller J.A. (1985). *Self-organization of predictive models*. Kiev: Technique; Berlin: FEB Verlag Technik, 1985. 223 p. (in Russian).]
- Малов Д.Н. (2019). Оценка инвестиционной привлекательности компаний на основе модели VAR (векторной авторегрессии) и ARIMA с учетом рисков // *Инновации и инвестиции*. № 1. С. 152–159. [Malov D.N. (2019). Assessing the investment attractiveness of companies based on the VAR (vector autoregression) and ARIMA models, taking into account risks. *Innovations and Investments*, no. 1, pp. 152–159 (in Russian).]
- Маматова Н. (2015). Применение модели векторной авторегрессии для анализа потребления электроэнергии // *Математические модели экономики: сборник научных трудов*. М.: НИУ ВШЭ. № 4. С. 15–19. [Mamatova N. (2015). Application of a vector autoregression model for the analysis of electricity consumption. *Mathematical Models of Economics: Collection of Scientific Papers* (NRU HSE), no. 4, pp. 15–19. (in Russian).]
- Мехович С.А., Ахизер Е.Б., Дунаевская О.И. (2014). Экономико-математическая модель зонирования промышленных предприятий // *Энергосбережение, энергетика, энергоаудит*. № 8 (126). С. 39–49. [Mekhovich S.A., Akhiezer E.B., Dunaevskaya O.I. (2014). Economic and mathematical model of zoning of industrial enterprises. *Energy Saving, Power Engineering, Energy Audit*, no. 8 (126), pp. 39–49 (in Russian).]
- Петров К.Э., Дейнеко А.А., Чалая О.В., Панферова И.Ю. (2020). Метод ранжирования альтернатив при проведении процедуры коллективного экспертного оценивания // *Радиоэлектроника, информатика, управления*. № 2. С. 84–94. [Petrov K.E., Deineko A.A., Chalaya O.V., Panferova I.Yu. (2020). Method for ranking alternatives during the procedure of collective expert assessment. *Radioelectronics, Informatics, Management*, no. 2, pp. 84–94 (in Russian).]
- Погосян К. (2015). Альтернативные модели прогнозирования основных макроэкономических показателей в Армении // *Квантиль*. № 13. С. 25–39. [Poghosyan K. (2015). Alternative models for forecasting the main macroeconomic indicators in Armenia. *Quantile*, no. 13, pp. 25–39 (in Russian).]
- Салманов О.Н., Заернюк В.М., Лопатина О.А. (2016). Установление влияния денежно-кредитной политики методом векторной авторегрессии // *Финансы и кредит*. № 28. С. 2–17. [Salmanov O.N., Zaernyuk V.M., Lopatina O.A. (2016). Establishing the influence of monetary policy using the vector autoregression method. *Finance and Credit*, no. 28, pp. 2–17 (in Russian).]

- Светульников С.Г., Баженова М.П., Лукаш Е.В. (2022). Перспективы использования векторных авторегрессий в экономическом прогнозировании // Современная экономика: проблемы и решения. № 6 (150). С. 44–57. [Svetunkov S.G., Bazhenova M.P., Lukash E.V. (2022). Prospects for the use of vector autoregressions in economic forecasting. *Modern Economics: Problems and Solutions*, no. 6 (150), pp. 44–57 (in Russian).]
- Суханова Е.И., Ширнаева С.Ю. (2014). Прогнозирование показателей стабилизационных процессов экономики России на основе моделей векторной авторегрессии // Фундаментальные исследования. № 9. С. 1590–1595. [Sukhanova E.I., Shirnaeva S.Yu. (2014). Forecasting indicators of stabilization processes in the Russian economy based on vector autoregression models. *Fundamental Research*, no. 9, pp. 1590–1595 (in Russian).]
- Шимановский Д.В., Третьякова Е.А. (2020). Моделирование социо-эколого-экономических взаимосвязей как способ оценки устойчивого развития регионов РФ // Вестник Пермского университета. Серия: Экономика. № 3 (15). С. 369–384. [Shimanovsky D.V., Tretyakova E.A. (2020). Modeling socio-ecological-economic relationships as a way to assess the sustainable development of regions of the Russian Federation. *Perm University Herald. Series: Economics*, no. 3 (15), pp. 369–384 (in Russian).] DOI: 10.17072/1994-9960-2020-3-369-384
- Ahelegbey D.F., Billio M., Casarin R. (2016). Special issue on recent developments in financial econometrics. *Annals of Economics and Statistics*, no. 123/124, pp. 333–361. DOI: 10.15609/annae-constat2009.123-124.0333
- Barrett A. (2021). Forecasting the Prices of Cryptocurrencies using a Novel Parameter Optimization of VARIMA Models. Chapman: Chapman University Digital Commons. 277 p.
- Carolyn N.N., Sherris M. (2020). Modeling mortality with a Bayesian vector autoregression. *Mathematics and Economics*, no. 94, pp. 40–57.
- Chandra S.R., Al-Deek H. (2009). Predictions of Freeway Traffic Speeds and Volumes Using Vector Autoregressive Models. *Journal of Intelligent Transportation Systems*, vol. 13, no. 2, pp. 53–72.
- Fitrianti H., Belwawin S.M., Riyana M., Amin R. (2019). Climate modeling using vector moving average autoregressive. *IOP Conference Series: Earth and Environmental Science*. Jaipur, pp. 28–39.
- Garcia-Martos C., Rodriguez J., Sanchez M.J. (2013). Modelling and forecasting fossil fuels, CO2 and electricity prices and their volatilities. *Applied Energy*, vol. 101, pp. 363–375.
- Jusmawati M.H., Penerapan V. (2020). Model Vector Autoregressive Integrated Moving Average dalam Peramalan Laju Inflasi dan Suku Bunga di Indonesia. *EIGEN mathematics journal*, 2020, december, no. 3 (2), pp. 73–82.
- Kilian L., Lütkepohl H. (2017). Structural Vector Autoregressive Analysis. Cambridge: Cambridge University Press. 735 p. DOI: <https://doi.org/10.1017/9781108164818>
- Lusia D.A., Ambarwati A. (2018). Multivariate Forecasting Using Hybrid VARIMA-Neural Network in JCI Case, Proceeding // International Symposium on Advanced Intelligent Informatics: Revolutionize Intelligent Informatics Spectrum for Humanity. Yogyakarta. Pp. 11–14.
- Lütkepohl H. (2005). New Introduction to Multiple Time Series Analysis. Berlin: Springer. 764 p.
- Meimela A., Lestari S.S. et al. (2021). Modeling of COVID-19 in Indonesia using vector autoregressive integrated moving average. *Journal of Physics: Conference Series*, pp. 55–79.
- Olson D.R., Riedel T.O. et al. (2021). Time series analysis of wintertime O3 and NOx formation using vector autoregressions. *Atmospheric Environment*, vol. 259, pp. 218–232.
- Ord K., Fildes R., Kourentzes N. (2017). Principles of business forecasting. Wessex: Wessex Press, Inc. 544 p.
- Rusyana A., Tatsara N., Balqis R., Rahmi S. (2020). Application of Clustering and VARIMA for Rainfall Prediction. *IOP Conference Series: Materials Science and Engineering*, pp. 428–438.
- Schmidhuber J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, no. 61, pp. 85–117.
- Thelin E.P., Raj R. et al. (2020). Comparison of high versus low frequency cerebral physiology for cerebrovascular reactivity assessment in traumatic brain injury: a multi-center pilot study. *Journal of*

- Clinical Monitoring and Computing*, no. 34 (5), pp. 971–994.
- Yi Zh., Chuntian Ch. et al. (2021). Multivariate probabilistic forecasting and its performance's impacts on long-term dispatch of hydro-wind hybrid systems. *Applied Energy*, vol. 283. pp. 116–243.
- Xu Bin, Boqiang Lin. (2016). What cause a surge in China's CO₂ emissions? A dynamic vector autoregression analysis. *Journal of Cleaner Production*, vol. 143, pp. 17–26.
- Zeiler F.A., Ercole A., Cabeleira M. (2020). Evaluation of the relationship between slow-waves of intracranial pressure, mean arterial pressure and brain tissue oxygen in TBI: a CENTER-TBI exploratory analysis. *Journal of Clinical Monitoring and Computing*, no. 35 (4), pp. 781–799.
- Zhang Cheng, Liao Huchang, Luo Li, Xu Zeshui (2021). Low-carbon tourism destination selection by a thermodynamic feature-based method. *Journal of the Operational Research Society*, June.

Рукопись поступила в редакцию 07.05.2024

METHOD FOR CONSTRUCTING VECTOR AUTOREGRESSIONS OF ANY COMPLEXITY

S.G. Svetunkov

DOI: 10.33293/1609-1442-2024-3(106)-37-50

EDN: LXPIGW

Sergey G. Svetunkov, Doct. Sc. (Economics), Professor at the Higher School of Business Engineering Peter the Great, St. Petersburg Polytechnic University, St. Petersburg, Russia; sergey@svetunkov.ru; eLibrary SPIN: 1772-7180

Acknowledgement. The work was carried out with the financial support of the Russian Science Foundation, grant No. 19-010-00610/19 “Theory, methods and techniques for forecasting economic development using autoregressive models of complex variables”.

Abstract. Vector autoregressions are one of the rapidly developing areas of many areas of modern science. They are actively used in modeling and forecasting various economic processes, most often in modeling the stock market and retail prices. Their most important advantage is the ability to consider the simultaneous influence of modeled indicators not only from their past values, but also from the past values of other indicators interrelated with them. The main problem why vector autoregressions are not actively used in practice (as they deserve) is the “curse of dimensionality”, which consists in a quadratic increase in the number of model coefficients depending on the increase in the dimension of the modeled vector. This circumstance leads to the fact that researchers in various fields of modern science are forced to limit the dimension of the vector, including only the most important ones in the model, or reducing the order of autoregression. Attempts to overcome the “curse of dimensionality” by using special mathematical methods result in a significant complication of the mathematical apparatus for constructing vector autoregressions, which does not contribute to the expansion of the practice of using vector autoregressions. The article proposes to use for this purpose a step-by-step decomposition method for constructing vector autoregressions of any dimension, which makes the process of constructing these models simple and accessible to any researcher. To test the possibility of using this method in practice, data series on the dynamics of eight main industry indices of the Moscow Exchange were used. At the same time, it was decided to construct a large vector autoregression of the order of $p = 10$. Using the least squares method, a total of 648 unknown coefficients of this model were estimated. The verification of this model was confirmed by simple autoregressions.

Keywords: vector autoregression, least squares method, dimension, lag, step-by-step decomposition method.

JEL classification: C22, C29, C53.

For reference: Svetunkov S.G. (2024). Method for constructing vector autoregressions of any complexity. *Economics of Contemporary Russia*, no. 3 (106). Pp. 37–50 (in Russian). DOI: 10.33293/1609-1442-2024-3(106)-37-50; EDN: LXPIGW

Manuscript received 07.05.2024