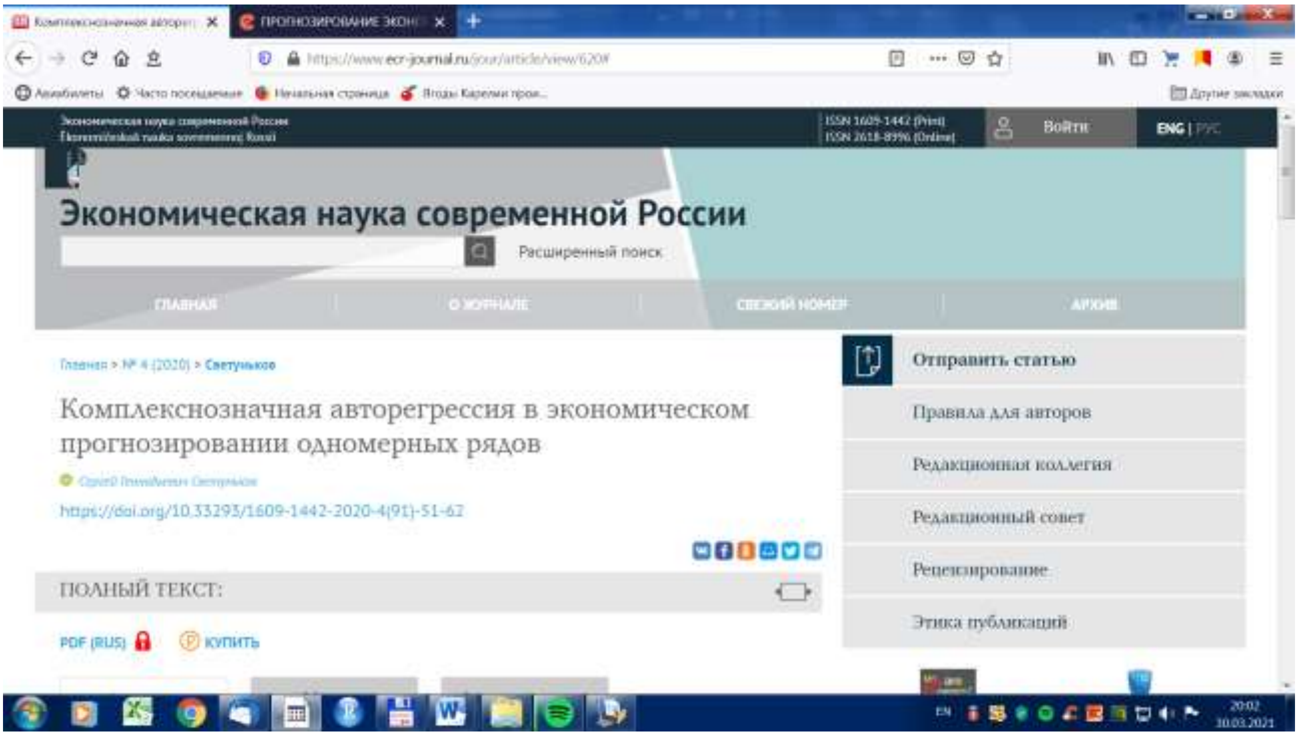

Светуньков С.Г. Комплекснозначная авторегрессия в экономическом прогнозировании одномерных рядов. *Экономическая наука современной России*. 2020;(4):51-62.
https://doi.org/10.33293/1609-1442-2020-4(91)-51-62

Аннотация

При краткосрочном прогнозировании экономических процессов активно используют модели авторегрессии *AR(p)* и их многочисленные модификации. При этом не всегда удаётся добиться необходимой точности прогноза, поэтому учёные, занимающиеся экономическим прогнозированием, продолжают разрабатывать новые методы и подходы для того, чтобы с их помощью повысить точность своих прогнозов. Один из перспективных подходов в этом направлении связан с использованием элементов теории функций комплексной переменной в моделировании экономики (комплекснозначная экономика). В статье показано, как, используя комплекснозначные авторегрессионные модели, повысить точность краткосрочного экономического прогнозирования. Рассматриваются свойства и возможность практического применения в краткосрочном экономическом прогнозировании двух моделей: модели комплекснозначной авторегрессии, к действительной части которой относится прогнозируемый показатель, а к мнимой – время, в которое этот показатель наблюдался (модель *CTAR(p)),* и модели, к действительной части которой относится прогнозируемый показатель, а к мнимой части – текущая ошибка прогноза *CARE(p)*. Показывается, что классическая модель авторегрессии действительных переменных *AR(p)* является частным случаем каждой из этих двух моделей. Основной акцент в статье делается на изучении свойств модели Re*CARE(p)*. Теоретически обосновывается, что эта модель точнее прогнозирует краткосрочную экономическую динамику, нежели модель *AR(p)*. На практических примерах показывается это. Поэтому рекомендуется в тех случаях, в которых уместны модели авторегрессии, использовать новую модель Re*CARE(p)*, как более точную. Показывается, что на основе этой базовой модели можно разработать новые модели краткосрочного прогнозирования, аналогичные моделям *ARMA(p,q)* и *ARIMA(p,d,q)*, от которых следует ожидать повышенной точности краткосрочных экономических моделей.

Ключевые слова. Комплекснозначная экономика, краткосрочное экономическое прогнозирование, авторегрессии.

JEL classification: C22, C29, C53

ВВЕДЕНИЕ

При краткосрочном прогнозировании динамики различных экономических процессов довольно часто используются модели авторегрессии (*AR*) (Corba, 2020; Fildes, 2020). В самом простом случае эта модель может быть записана так:

$$y_{t+1} = by_t . \quad (1)$$

Свойства этой модели давно изучены и потому известно, что в динамике эта модель описывает траекторию степенной функции:

$$y_{t+1} = b^t y_0 . \quad (2)$$

В том случае, когда коэффициент авторегрессии *b* больше единицы, модель в долгосрочной перспективе генерирует экспоненциальный рост. В том случае, когда коэффициент авторегрессии меньше единицы, модель *AR* генерирует в долгосрочной перспективе убывающую последовательность значений. Но для краткосрочного экономического прогнозирования эта модель оказывается удобной в использовании и довольно точной для прогнозирования некоторых экономических рядов.

Модель авторегрессии первого порядка (1), проста в использовании, а потому нашла широкое применение в практике экономического прогнозирования (Ord, 2017). Более сложными, а потому не столь часто используемыми являются модели порядка *p*, которые обозначают так: *AR(p)*. Для определения лага и соответственно порядка авторегрессии можно построить автокорреляционную функцию, хотя в последние годы для этого используют подход, называемый «информационный критерий» - с помощью компьютерной программы модель *AR(p)* последовательно меняет свой лаг *p*, и для каждого лага вычисляется дисперсия ошибки ретропрогноза и информационный критерий, базирующийся на этой дисперсии. Выбирается тот лаг, для которого величина информационного критерия минимальна (Vu Ky, 2007, p.213).

Прогнозное значение показателя определяется как его расчётное значение плюс ошибка аппроксимации. Применительно к модели *AR(1)* это может быть записано так::

$$y_{t+1} = by_t + \varepsilon_{t+1} . \quad (3)$$

Но если прогнозируется первое слагаемое равенства (3), то почему бы не попытаться спрогнозировать его второе слагаемое, а именно – ошибку аппроксимации? Предполагая о том, что ошибки аппроксимации гомоскедастичны, в качестве прогнозной модели ошибок аппроксимации можно было бы использовать их среднюю арифметическую. Но чаще всего коэффициент авторегрессии *b* оценивают методом наименьших квадратов (МНК). А в этом случае сумма ошибок отклонений ε_t фактических значений от расчётных будет равна нулю:

$$\sum_{t=0}^{T} \varepsilon_t = 0 \quad (4)$$

Поэтому среднюю арифметическую использовать нельзя, а вот если вычислять среднюю арифметическую нескольких последних членов ряда, то она в общем случае не равна нулю:

$$\bar{\varepsilon}_q = \frac{1}{q} \sum_{t=T-q}^{T} \varepsilon_t \neq 0$$

и уже эту среднюю арифметическую можно использовать для прогнозирования ошибки аппроксимации ε_{t+1}.

Давно уже доказано, что вместо простой скользящей средней для экономического прогнозирования лучше использовать взвешенную среднюю:

$$\bar{\varepsilon}_q = \sum_{t=T-q}^{T} c_t \varepsilon_t \neq 0, \quad \sum_{t=T-q}^{T} c_t = 1 \quad (5)$$

Такую скользящую взвешенную среднюю принято обозначать *MA(q)*, где *q* - период усреднения. Модель авторегрессии с прогнозируемыми ошибками прогноза принято обозначать *ARMA(p,q)*.

В последние годы требование того, чтобы сумма весовых коэффициентов в модели *MA(q)* была равна единице, уже не выступает таким жёстким, как изначально, и потому включение в название модели абревиатуры *MA* (*moving average*) сегодня следует рассматривать как атавизм, а не как обязательное требование. Но вышеизложенная логика формирования модели *ARMA(p,q)* в целом сохранилась.

Эта модель в современной прогностике является наиболее популярной, поскольку она, помимо уже указанных преимуществ, включает в себя как частный случай экспоненциальное сглаживание - другой, не менее популярный метод краткосрочного прогнозирования в экономике (Лукашин, 2017).

И модель *AR(p)*, и модель *ARMA(p,q)* в своей основе имеют траекторию степенного тренда, что ограничивает сферу их успешного применения на практике. Поэтому современные экономисты, занимающиеся экономическим прогнозированием, ищут пути улучшения точности модели *ARMA(p,q)* с помощью различных модификаций этой модели. В частности, стремясь избавиться от нелинейного роста, в модель подставляют не исходные данные, а их приращения, и такая модель получила название *ARIMA(p,d,q)*, где *d* – порядок разности. Комплекс этих моделей и методов краткосрочного прогнозирования, не смотря на свою разработанность и опыт успешного практического применения, не всегда справляется с задачей краткосрочного прогнозирования экономики, а потому учёные продолжают разрабатывать различные модификации этих моделей (Gully, 2019).

В том случае, если прогнозируются многомерные ряды, учёные начали применить модели векторной авторегрессии *VAR*, однако, вычислительные сложности ограничивают их применение (Vu Ky, 2007). Частным случаем моделей *VAR* являются комплекснозначные авторегрессионные модели *CAR*, которые имеют в два раза меньше коэффициентов, чем модели *VAR*, а потому могут быть особенно эффективными при прогнозировании коротких рядов.

ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ

Одним из перспективных направлений экономико-математического моделирования, и экономического прогнозирования в том числе, является комплекснозначная экономика, в которой в качестве основных переменных используются комплексные переменные. Если в модель авторегрессии (1) вместо действительной переменной подставить комплексную переменную ($y_{rt}+iy_{it}$), а вместо действительного коэффициента использовать комплексный коэффициент авторегрессии (a_0+ia_1), то будет получена комплекснозначная авторегрессионная модель (*CAR*). Для авторегрессии первого порядка она будет выглядеть так (Svetunkov, 2012, с. 297):

$$y_{r(t+1)} + iy_{i(t+1)} = (a_0 + ia_1)(y_{rt} + iy_{it}). \quad (6)$$

Здесь y_r и y_i – прогнозируемые переменные, a_0 и a_1 – коэффициенты *CAR*, *i* – мнимая единица ($i^2=-1$).

Свойства модели (6) существенно отличаются от свойств модели (1). Модель также моделирует динамику степенной функции

$$y_{r(t+1)} + iy_{i(t+1)} = (a_0 + ia_1)^t (y_{r0} + iy_{i0}), \quad (7)$$

но, поскольку функция является комплекснозначной, то на плоскости (y_r; y_i) будет моделироваться спиралеобразная динамика изменения показателей y_r и y_i. При этом если модуль комплексного коэффициента авторегрессии (a_0+ia_1) будет больше единицы, то спираль раскручивается; если он меньше единицы, то спираль закручивается; если он равен единице, то описывается динамика по окружности, радиус которой равен модулю первоначального значения комплексной авторегрессии ($y_{r0}+iy_{i0}$).

При этом каждая составляющая модели (7), рассмотренная по отдельности в динамике, демонстрирует колебательную траекторию развития, соответственно либо с возрастающим размахом колебаний, либо с затухающим размахом колебаний.

Для использования модели *CAR* в экономическом прогнозировании одномерных рядов, которые сегодня прогнозируются с помощью *ARMA(p,q),* необходимо адаптировать модель *CAR* для этого одномерного случая, то есть, сформировать такую комплексную переменную, в действительной части которой будет находиться прогнозируемый показатель, а в мнимой часть – его некоторая дополнительная характеристика прогнозируемого показателя.

ДВЕ МОДЕЛИ ОДНОМЕРНЫХ КОМПЛЕКСНЫХ АВТОРЕГРЕССИЙ

Рассмотрим варианты адаптации модели *CAR* к прогнозированию одномерных рядов.

Здесь и далее будем рассматривать самый простой случай, когда порядок авторегрессии равен единице, понимая, что переход к более сложным моделям не окажется затруднительным. При необходимости мы этот переход и будем осуществлять.

Первым вариантом адаптации модели *CAR* к одномерному ряду будет являться комплексная авторегрессионная модель, в которой одной из факторных, а, следовательно, и вычисляемых переменных, является время t (Светуньков, 2020):

$$y_{t+1}+i(t+1)=(a_0+ia_1)(y_t+it). \qquad (8)$$

Назовём эту модель, как это принято в современной экономической прогностике, по заглавным буквам – *CTAR*. Прежде всего, обратим внимание на то, что в модели (8) прогнозируется не только экономический показатель y_{t+1}, но и время ($t+1$). В моделях авторегрессии время выступает как индекс упорядочивания показателей, но не как активный элемент, учитываемый в прогнозировании и прогнозируемый, а в (8) время выступает и как определяющий фактор, и как зависимая переменная.

Поскольку любая комплексночначная функция может быть представима как система двух равенств – отдельно для вещественной и мнимой частей комплекснозначной функции, то получим для модели (8):

$$\begin{cases} y_{t+1}=a_0y_t-a_1t, \\ (t+1)=a_0t+a_1y_t. \end{cases} \qquad (9)$$

Первое уравнение системы (9) показывает, как вычисляется прогнозируемый показатель в зависимости от значений y_t и времени t. Второе уравнение показывает, что в модели *CTAR* вычисляемый прогнозный момент времени ($t+1$) рассчитывается как линейная двухфакторная регрессия в зависимости от значения y_t и времени t, когда этот показатель наблюдался. Это значит, что вторая составляющая комплексной переменной модели *CTAR* только в редких случаях будет целочисленной, а чаще всего она будет дробной. То есть, модель (8) прогнозирует нецелый период прогноза. Поэтому если с помощью модели (8), например, будет прогнозироваться такое комплексное число – *1345,7+i2021,3*, то моменту наблюдения 2021,3 можно, конечно, поставить в соответствие месяц и день 2021 года, но экономиста интересует результат не на конкретный день апреля 2021 года, а на конец этого года. Поэтому здесь необходимо использовать соответствующие процедуры «подтягивания» времени до целого значения, и пересчёта прогнозируемого показателя, соответствующего этому целому значению времени. Это неудобство ограничивает сферу применения модели *CTAR* в экономическом прогнозировании, хотя определённый интерес для прогнозирования экономики она представляет. Модель Re*CTAR* представляется заурядной.

Второй вариант адаптации модели *CAR* к задачам прогнозирования одномерных рядов является более перспективным. Ещё в самом начале формирования комплекснозначной экономики И.С. Светуньков предложил модификацию метода экспоненциального сглаживания, в которой вместо действительной переменной он использовал комплексную переменную такого вида: $(y_t + i(y_t - \hat{y}_t))$. Здесь, как видно, в комплексную переменную включаются прогнозируемый показатель и текущее значение ошибки аппроксимации, что приводит к появлению у модели экспоненциального сглаживания новых свойств (Svetunkov I., Kourentzes N., 2015). Воспользовавшись этой идеей, представим комплексную авторегрессию *CAR* в таком виде:

$$y_{t+1} + i(y_{t+1} - \hat{y}_{t+1}) = (a_0 + ia_1)(y_t + i(y_t - \hat{y}_t)). \quad (10)$$

Этой модели, которую будем называть *CARE*, в области действительных переменных соответствует такая система равенств:

$$\begin{cases} y_{t+1} = a_0 y_t - a_1(y_t - \hat{y}_t), \\ y_{t+1} - \hat{y}_{t+1} = a_0(y_t - \hat{y}_t) + a_1 y_t. \end{cases} \quad (11)$$

И, как видно из этой системы, прогнозное значение y_{t+1} «корректируется» на ошибку $(y_t - \hat{y}_t)$, а прогнозная величина ошибки аппроксимации $(y_{t+1} - \hat{y}_{t+1})$ прогнозируется с помощью текущего значения показателя y_t и текущего значения ошибки аппроксимации.

Эта модель свободна от недостатка модели *CTAR*, когда необходимо давать некоторую интерпретацию дробному времени, и, более того, она обладает целым рядом неоспоримых достоинств, на которые и следует обратить внимание.

Прежде всего, необходимо отметить, что том случае, когда в комплексном коэффициенте авторегрессии (9) равен нулю коэффициент мнимой части ($a_1=0$), то модель *CARE(1)* превращается в модель *AR(1)*, поскольку прогнозируемая действительная часть модели (10) в таком случае будет равна:

$$y_{t+1} = a_0 y_t, \quad (12)$$

то есть, получается действительная авторегрессия первого порядка.

Во всех остальных случаях модель *CARE* будет отличаться от модели авторегрессии действительных переменных *AR*. Это означает, что модель *CARE* описывает более широкий класс моделей авторегрессий действительных переменных, нежели модель *AR*. Последняя, как видно, является частным случаем *CARE*.

Для того чтобы быть уверенным в пригодности модели *CARE* в экономическом прогнозировании, необходимо выявить её основные свойства.

Вспомним, что комплекснозначная авторегрессионная модель описывает динамику, соответствующую степенной комплекснозначной функции, то есть модель (10) можно записать и так:

$$y_{t+1} + i(y_{t+1} - \hat{y}_{t+1}) = (a_0 + ia_1)(y_t + i(y_t - \hat{y}_t)) = (a_0 + ia_1)^t (y_0 + i(y_0 - \hat{y}_0)). \quad (13)$$

Запишем комплексный коэффициент пропорциональности и начальное значение комплексной переменной в экспоненциальной форме:

$$a_0 + ia_1 = \sqrt{a_0^2 + a_1^2}\, e^{iarctg\frac{a_1}{a_0}} = Re^{i\theta}, \quad y_0 + i(y_0 - \hat{y}_0) = R_y e^{i\varphi}. \quad (14)$$

С учётом этого модель *CARE(p)* может быть представлена так:

$$y_{t+1} + i(y_{t+1} - \hat{y}_{t+1}) = R^t e^{it\theta} R_y e^{i\varphi} = R^t R_y e^{i(t\theta+\varphi)}. \quad (15)$$

Поскольку нас при прогнозировании одномерного ряда интересует действительная часть авторегрессии *CARE*, то есть *ReCARE*, то из (15) следует:

$$y_{t+1} = \mathrm{Re}[R^t R_y e^{i(t\theta+\varphi)}] = R^t R_y \cos(t\theta + \varphi). \quad (16)$$

Тогда понятно, как ведёт себя эта модель.

Если, например, модуль коэффициента пропорциональности R больше единицы, а его полярный угол θ положителен, то прогнозируемый с помощью *CARE(p)* экономический показатель y_t имеет сложную нелинейную динамику. В этом случае с ростом времени t он увеличивается в R^t раз и одновременно уменьшается, поскольку при положительности полярного угла θ с ростом времени угол ($t\theta+\varphi$) также растёт, а косинус этого возрастающего в первом квадранте угла уменьшается до нуля, а в долгосрочной перспективе он и вовсе становится отрицательным. Но поскольку мы рассматриваем задачи краткосрочного прогнозирования, то долгосрочная перспектива не представляет особого интереса.

В том случае, когда модуль комплексного коэффициента пропорциональности меньше единицы, то R^t, как первый сомножитель модели *CARE*, с ростом времени уменьшается. Если при этом полярный угол комплексного коэффициента пропорциональности θ находится в пределах ($3/2\pi \geq \theta \geq 2\pi$), то с ростом t сомножитель $\cos(t\theta+\varphi)$ увеличивается. В этом случае моделируется другая динамика роста прогнозируемого показателя, также нелинейная, но отличная от той, которая была рассмотрена ранее.

То есть, прогнозируемая действительная часть модели *CARE(1)* может иметь самую разную нелинейную тенденцию в зависимости от значений комплексного коэффициента пропорциональности. Более того, при определённом сочетании значений комплексного коэффициента пропорциональности и времени модель Re*CARE(1)* будет моделировать линейную динамику прогнозируемой действительной части, что для модели *AR(p)* невозможно – модель *AR(p)* линейна только в одном случае: когда $b=1$. В этом случае прямая линия будет параллельна оси времени и характеризоваться условием $y_t=y_0$.

Поэтому предлагаемая модель Re*CARE* обладает более разнообразными свойствами, нежели модель *AR* и способна успешно прогнозировать такую экономическую динамику, которая модель *AR* прогнозирует плохо.

ИССЛЕДОВАНИЕ СВОЙСТВ МОДЕЛИ Re*CARE*(1)

Для того, чтобы получить дополнительные основания и быть уверенными в том, что модель Re*CARE(1)* может с успехом использоваться в краткосрочном экономическом прогнозировании вместо модели *AR(1)*, следует проверить её свойства на практических примерах и сравнить точность предлагаемой модели Re*CARE(1)* с точностью прогнозирования классической модели авторегрессии *AR(1)*. Для этого на фактических статистических данных необходимо оценить коэффициенты модели (10).

Прежде всего, запишем модель в форме регрессионной зависимости.

Фактическая пара будущих значений $y_{t+1}-i(y_{t+1}-\hat{y}_{t-1})$ рассчитывается через предшествующую пару фактических значений $y_t - i(y_t - \hat{y}_t)$ с некоторой ошибкой аппроксимации($\varepsilon_{r(t+1)}+i\varepsilon_{i(t+1)}$), которая является комплексной случайной величиной. С учётом этих обозначений уравнение регрессии *CARE(1)* можно записать так:

$$y_{t+1}+i(y_{t+1}-\hat{y}_{t+1})=(a_0+ia_1)(y_t+i(y_t-\hat{y}_t))+(\varepsilon_{r(t+1)}+i\varepsilon_{i(t+1)}) . \quad (17)$$

Для практического использования модели следует найти коэффициенты комплексной авторегрессии (a_0+ia_1), минимизируя сумму квадратов комплексных отклонений

$$\sum(\varepsilon_{r(t+1)}+i\varepsilon_{i(t+1)})^2 \rightarrow \min . \quad (18)$$

Соответствующий метод и методики МНК изложены в (Светуньков, 2019). Поэтому здесь на этом вопросе останавливаться не будем. Но поскольку нас интересует прогноз не всей комплексной переменной, а только её действительной части (y_{t+1}), то задача становится ординарной. Из (17) следует, что Re*CARE(p)* может быть представлена в таком виде:

$$y_{t+1}=a_0y_t-a_1(y_t-\hat{y}_t)+\varepsilon_{r(t+1)} . \quad (19)$$

Мнимая составляющая, то есть модель Im*CARE* в данном случае не представляет интереса, а для модели (19) нахождение коэффициентов a_0 и a_1 с помощью метода наименьших квадратов – тривиальная задача.

Проверим на практических примерах прогнозную точность модели Re*CARE* и сравним её с моделью *AR*. Будем использовать для этого простую процедуру. Представим и модель (19), и модель (1) в виде моделей, прогнозирующих показатель на τ наблюдений вперёд, где $\tau=1,2,..,10$:

$$y_{t+\tau} = a_0 y_t - a_1(y_t - \hat{y}_t) + \varepsilon_{r(t+\tau)}, \qquad (20)$$

$$y_{t+\tau} = b y_t + \varepsilon_{t+\tau}. \qquad (21)$$

Сначала сравним свойства моделей (20) и (21) на примере прогноза количества ежедневно фиксируемых в мире заболевших коронавирусом COVID19. Используя статистические данные по этому заболеванию со 2 мая по 7 июля 2020 года (67 наблюдений), проведём процедуру сравнения точности этих двух авторегрессий. С помощью МНК будем находить коэффициенты этих двух моделей для различных прогнозных периодов τ, которые последовательно меняются от 1 до 10 и вычислять для каждого τ среднеквадратичное отклонение (СКО) ошибок ретропрогноза. Результаты вычислений представлены в таблице 1.

Таблица 1.

Сравнение моделей Re*CARE и AR на примере ежедневной статистики заболевания COVID 19 в мире*

Прогнозный период, τ	a_0	a_1	b	*СКО для модели* Re*CARE*	*СКО для модели AR*	*% уменьшения СКО*	*% отличия a_0 от b*
1	1,016	0,943	0,978	3572,766	5074,065	34,7	3,8
2	1,029	0,926	0,997	3537,808	4908,558	32,5	3,2
3	1,043	0,883	1,035	3525,822	4754,041	29,7	0,8
4	1,057	0,939	1,037	3534,535	4792,845	30,2	2,0
5	1,069	0,907	1,054	3330,863	4354,825	26,6	1,4
6	1,082	0,557	1,070	3594,825	3906,339	8,3	1,2
7	1,091	0,983	1,081	3277,831	4024,852	20,5	0,9
8	1,105	0,931	1,102	3603,864	3973,140	9,7	0,3
9	1,119	0,885	1,109	3757,969	4621,566	20,6	0,9
10	1,133	0,976	1,117	3446,661	4690,614	30,6	1,4

К этой таблице необходимо дать некоторые пояснения.

В первом столбце таблицы приводится величина прогнозного периода авторегрессии, во втором и в третьем - коэффициенты модели Re*CARE* и модели *AR*, в четвёртом и пятом столбцах - значения соответствующих каждой из рассматриваемых моделей среднеквадратичных отклонений ошибок ретропрогноза. Предпоследний столбец этой таблицы показывает, на сколько процентов ретропрогноз с помощью модели Re*CARE* лучше описывает этот ряд, чем модель *AR*. Как и ожидалось, повышение точности наблюдается во всех рассмотренных случаях.

По последнему столбцу таблицы 1 видно, что коэффициент b довольно близок по своим значениям к коэффициенту a_0. А для лага в восемь наблюдений эти два коэффициента отличаются друг от друга всего лишь на 0,3%, то есть – они почти равны друг другу. Но в этом случае коэффициент a_1 не равен нулю, что и обеспечивает модели Re*CARE* большую точность, чем модели *AR* на 9,7%. Такое повышение точности прогноза следует оценить как весомое.

Интересно, что между двумя последними столбцами таблицы 1, то есть – между процентом повышения точности модели Re*CARE* по сравнению с моделью *AR* и процентом отличия коэффициента a_0 от коэффициента b, имеется довольно высокий коэффициент парной корреляции, который равен 0,68. Это означает, что в целом наблюдается такая линейная тенденция – чем сильнее коэффициенты этих моделей отличаются друг от друга, тем точнее модель Re*CARE*. Или, иначе говоря: чем хуже прогнозирует классическая модель *AR*, тем лучше прогнозирует модель Re*CARE*.

В своё время С. Макридакис сформировал общедоступную базу данных самых разных динамических рядов экономических процессов, которой сегодня пользуются многие учёные, занимающиеся экономическим прогнозированием. На этой базе они проверяют свойства новых прогнозных моделей и сравнивают их точность друг с другом. Интересные результаты были получены по ряду №2830 этой базы при сравнении прогностических свойств моделей Re*CARE* и *AR*. Как было показано в теоретической части статьи, модель *AR* является частным случаем модели Re*CARE*. И модель Re*CARE* превращается в модель *AR* только в том случае, когда коэффициент a_1 равен нулю, а $a_0=b$. Этот случай и был зафиксирован применительно к этому ряду базы С. Макридакиса в одном из вариантов ретропрогноза. Результаты сравнительного анализа, осуществлённого по точно такой же процедуре, как и раньше, приведены в таблице 2.

Таблица 2.

Сравнение моделей Re*CARE и AR на примере ряда №2830 базы С.Макридакиса*

Прогнозный период, τ	a_0	a_1	b	*СКО для модели* Re*CARE*	*СКО для модели AR*	*% уменьшения СКО*	*% отличия a_0 от b*
1	1,002	-0,251	1,000	12,414	12,829	3,3	0,2
2	1,004	0,005	1,004	20,274	20,274	0,0	0,0
3	1,007	0,134	1,006	25,176	25,348	0,7	0,1
4	1,010	0,210	1,008	28,655	29,045	1,3	0,2
5	1,019	0,937	1,011	30,885	32,405	4,8	0,8
6	1,023	0,964	1,014	30,826	34,603	11,5	0,9
7	1,028	0,974	1,017	30,769	36,512	17,1	1,1
8	1,032	0,927	1,020	31,872	38,272	18,2	1,1
9	1,037	0,981	1,024	30,945	39,718	24,8	1,2
10	1,040	0,925	1,028	32,184	41,532	25,4	1,2

Из данных таблицы 2 можно заметить, что для лага, равного двум, и модель Re*CARE*, и модель *AR* обладают одинаковой точностью ретропрогноза и при этом коэффициент a_1 практически равен нулю, а $a_0=b=1,004$.

Во всех остальных случаях для других лагов модель Re*CARE* всегда точнее модели *AR*. И вновь можно заметить взаимосвязь между процентом повышения точности модели Re*CARE* по сравнению с моделью *AR* и процентом отличия коэффициента a_0 от коэффициента b. При этом коэффициент парной корреляции между этими двумя столбцами равен 0,93. То есть и здесь наблюдается та же самая закономерность - чем сильнее коэффициенты этих моделей a_0 и b отличаются друг от друга, тем точнее модель Re*CARE* по сравнению с моделью *AR*.

СРАВНЕНИЕ МОДЕЛИ Re*CARE* И *ARMA*

Обратим внимание на внешнее сходство модели Re*CARE* с моделью *ARIMA*.

Построим вначале модель Re*CARE(p)*. Она будет иметь такой вид:

$$\hat{y}_{t+1} = \sum_{t=T-(p-1)}^{T} a_{0t} y_t - \sum_{t=T-(p-1)}^{T} a_{1t}(y_t - \hat{y}_t) \quad . \qquad (22)$$

Аналогом этой модели выступает модель *ARMA(p,q)*, которая с учётом принятых ранее обозначений запишется так:

$$\hat{y}_{t+1} = \sum_{t=T-(p-1)}^{T} b_t y_t - \sum_{t=T-(q-1)}^{T} c_{1t}(y_t - \hat{y}_t) \, . \qquad (23)$$

Поскольку требование равенства единице суммы коэффициентов составляющей MA(p) $\sum_{t=T-(p-1)}^{T} c_t = 1$ в настоящее время снято, то может показаться, что (22) и (23) равны друг другу в том случае, когда $q=p$. В этом случае и коэффициенты двух моделей будут равны друг другу, то есть $a_{ot}=b_t$, $a_{1t}=c_t$. Следует ли из этого, что модель Re*CARE(p)* представляется частным случаем модели *ARMA(p,q)*? На самом деле это не так. Схожесть моделей является мнимой.

Модели Re*CARE(p)* представляет собой целостную модель, коэффициенты которой можно находить на имеющейся базе данных, используя процедуру ретропрогноза, минимизируя сумму квадратов отклонений

$$\sum_t \varepsilon_{t+1}^2 = \sum_t (y_{t+1} - \hat{y}_{t+1})^2 \qquad (24)$$

или минимизируя базирующийся на СКО более сложный информационный критерий (Ord, 2017).

Модель *ARMA(p,q)* представляет собой аддитивную форму двух моделей – *AR(p)* и *MA(q)*, поэтому структурно она может быть записана так:

$$ARMA(p,q) = AR(p) + MA(q) \, . \qquad (25)$$

Эта форма записи предопределяет другую процедуру нахождения коэффициентов модели. Ещё в 1974 году Бокс и Дженкинс показали, что процессы *AR(p)* и *MA(q)* связаны друг с другом (Box, 2015). При этом авторегрессии конечного порядка *AR(p)* соответствует процесс скользящей средней бесконечного порядка *MA(∞)* и наоборот. Эта особенность и определяет процедуру нахождения коэффициентов данной модели (Zhang, 2018, p.324).

На практике, когда экономист, занимающийся прогнозированием, не обладает дополнительными математическими и вычислительными компетенциями, это делается так: вначале находятся коэффициенты b_t модели *AR(p)*; затем вычисляются ошибки аппроксимации, по которым вычисляются коэффициенты c_t модели *MA(q)*. Различие коэффициентов модели ReCARE и модели ARMA в этом случае очевидно.

В том случае, когда экономический прогноз выполняется с применением высококвалифицированных исполнителей, а точности самого прогноза предъявляются повышенные требования, используют более сложную процедуру. Поскольку модели *AR(p)* и *MA(q)* являются взаимосвязанными, то оценка коэффициентов b_t и c_t этих взаимосвязанных моделей может осуществляться с помощью многоитеративной процедуры поиска оптимальных значений этих коэффициентов и оптимальных значений *p* и *q* на основе суммы квадратов отклонений (24) или информационного критерия.

Последний третий случай нахождения коэффициентов модели *ARMA(p,q)* возможен для высококвалифицированных специалистов. Это – одновременное нахождение коэффициентов b_t и c_t этих моделей.

Как видно из анализа процесса построения модели *ARMA(p,q)*, возможности дальнейшего уточнения этой модели и повышения её прогнозной точности полностью исчерпаны.

А вот модель Re*CARE(p)* использует ограниченную часть ошибок аппроксимации, число которых соответствует порядку авторегрессии *p*. Поэтому она может быть дополнена и уточнена с помощью скользящей средней *MA(q)*, что структурно может быть записано так:

$$\mathrm{Re}\,CAREMA(p,q) = \mathrm{Re}\,CARE(p) + MA(q)\,. \qquad (26)$$

Из (22) следует, что *ReCARE(p) = ARMA(p,p)*. Подставляя это равенство в (26), получим:

$$\mathrm{Re}\,CAREMA(p,q) = ARMA(p,p) + MA(q)\,. \qquad (27)$$

То есть, модель Re*CAREMA(p,q)* и *ARMA(p,q)* отличаются друг от друга.

ЗАКЛЮЧЕНИЕ

Поскольку модель *CARE* впервые вводится в научный оборот, то в первой статье, описывающей её суть и её основные свойства, сложно рассказать обо всех возможностях и обо всех свойствах этой новой прогнозной модели. Но одно представляется неоспоримым: модель Re*CARE(p)* является более общей и более точной моделью, чем модель классической авторегрессии *AR(p)*, а потому при краткосрочном прогнозировании рекомендуется использовать именно её.

Дальнейшее исследование возможностей использования модели краткосрочного прогнозирования Re*CARE* связано со многими направлениями, в первую очередь с решением таких задач:

1. Необходимо более тщательно изучить свойства и возможности модели Re*CAREMA(p,q)* и Re*CAREIMA(p,d,q)*. Сравнить эти модели с *ARMA(p,q)* и *ARIMA(p,q)*.

2. В данной статье акцент сделан на исследование свойств действительной части модели *CARE(p)* – Re*CARE(p)*. Мнимая часть этой модели Im*CARE(p)*, а именно – некоторая оценка ошибки прогноза $(y_{t+1} - \hat{y}_{t+1})$ требует дополнительных исследований. Пока что не совсем понятен смысл этой вычисляемой ошибки.

3. Поскольку акцент в данном исследовании делался на действительной части комплекснозначной авторегрессии (10), то и коэффициенты этой модели рассматривались не как один комплексный коэффициент, а как два коэффициента двухфакторной модели действительных переменных (19). Они находились с помощью классического МНК. Если же рассматривать комплекснозначную авторегрессию (10) в целом, то необходимо использовать комплекснозначный МНК. В таком случае оценки коэффициентов будут отличаться от рассмотренного случая. Какими свойствами будет обладать действительная часть модели *CARE* при таком способе оценивания её коэффициентов, и как эти два способа оценивания коэффициентов взаимосвязаны друг с другом? Для ответа на эти вопросы следует провести дополнительные исследования.

Но в любом случае можно сделать вывод о том, что предлагаемая модель краткосрочного прогнозирования *CARE(p)* имеет научную значимость и её можно использовать в задачах краткосрочного экономического прогнозирования.

Список литературы

1. Лукашин Ю. П. Прогнозирование социально-экономических процессов: учебное пособие. Москва/Берлин: Директ-Медиа, 2017. 88 с.
2. Светуньков С.Г. Основы эконометрики комплексных переменных. СПб.: Медиапапир, 2019. 106 с.
3. Светуньков С.Г. Прогнозирование экономической динамики с помощью комплекснозначной авторегрессии с временной составляющей (CTAR) // Современная экономика: проблемы и решения. 2020, № 9. С. 21 – 31.
4. Box G.E. P., Jenkins G. M. Time series analysis, forecasting and control. John Wiley & Sons, May 29, 2015. 712 p.
5. Corba, B.S., Egrioglu, E., Dalar, A.Z. AR–ARCH Type Artificial Neural Network for Forecasting. Neural Processing Letters. 2020, N. 51, pp. 819–836.

6. Gully T. Non-Profit-Maximizing Behavior in Supply Chain Management. Springer, 2019. 171 p.
7. Fildes, R. Learning from forecasting competitions. International Journal of Forecasting, 2020, N. 36, pp. 3-18.
8. Ord K., Fildes R., Kourentzes N. Principles of Business Forecasting. Wessex, Incorporated, 2017. 588 p.
9. Svetunkov, I., Kourentzes, N. Complex Exponential Smoothing. Working Paper of Department of Management Science, Lancaster University 2015, 31 p.
10. Svetunkov Sergey. Complex-Valued Modeling in Economics and Finance. Springer Science+Business Media, New York, 2012. 318 p.
11. Vu Ky M. The ARIMA and VARIMA Time Series: Their Modelings, Analyses and Applications. AuLac Technologies Inc., 2007. 488 p.
12. Zhang Yi-xin, Sun Wen-sheng. Agricultural Product Price Forecast based on Shotr-term Time Series Analysis Techniques // Current Trends in Computer Science and Mechanical Automation Vol.1: Selected Papers from CSMA2016. Walter de Gruyter GmbH & Co KG, 2018. Pp. 317-328.

Complex-valued autoregression in economic forecasting of one-dimensional series

Sergey Svetunkov
Sergey Gennad'yevich Svetun'kov, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia, sergey@svetunkov.ru

Annotation
Autoregressive models *AR(p)* and their numerous modifications are actively used for short-term forecasting of economic processes. At the same time, it is not always possible to achieve the necessary forecast accuracy, therefore, scientists engaged in economic forecasting continue to develop new methods and approaches in order to use them to improve the accuracy of their forecasts. One of the promising approaches in this direction is associated with the use of elements of the theory of functions of a complex variable in modeling the economy (complex-valued economy). The article shows how, using complex-valued autoregressive models, to increase the accuracy of short-term economic forecasting. Here, the properties and the possibility of practical application in short-term economic forecasting of two models are considered: the complex-valued autoregression model, the real part of which is the predicted indicator, and the imaginary part is the time at which this indicator was observed (model *CTAR(p)*) and the model, the real part of which is the predicted indicator, and the imaginary part is the current forecast error *CARE(p)*. It is shown that the classical model of autoregression of real variables *AR(p)* is a special case of each of these two models. The main focus of the article is on studying the properties of the Re*CARE(p)* model. It is theoretically substantiated that this model predicts short-term economic dynamics more accurately than the *AR(p)* model. This is shown on practical examples. Therefore, it is recommended that in cases where autoregressive models are appropriate, use the new Re*CARE(p)* model, as it is more accurate. It is shown that, on the basis of this basic model, it is possible to develop new short-term forecasting models similar to the *ARMA(p,q)* and *ARIMA(p,q)* models, from which one should expect increased accuracy of short-term economic models.

Keywords: complex-valued economics, short-term economic forecasting, autoregressive models

References

1. Lukashin Yu. P. Forecasting socio-economic processes: a tutorial. Moscow / Berlin: Direct-Media, 2017. 88 p.
2. Svetunkov S.G. Fundamentals of Complex Variable Econometrics. SPb.: Mediapapir, 2019.106 p.
3. Svetunkov S.G. Forecasting economic dynamics using complex-valued autoregression with a time component (CTAR) // Modern economics: problems and solutions. 2020, No. 9.P. 21 - 31.
4. Box G.E. P., Jenkins G. M. Time series analysis, forecasting and control. John Wiley & Sons, May 29, 2015. 712 p.
5. Corba, B.S., Egrioglu, E., Dalar, A.Z. AR–ARCH Type Artificial Neural Network for Forecasting. Neural Processing Letters. 2020, N. 51, pp. 819–836.
6. Gully T. Non-Profit-Maximizing Behavior in Supply Chain Management. Springer, 2019. 171 p.
7. Fildes, R. Learning from forecasting competitions. International Journal of Forecasting, 2020, N. 36, pp. 3-18.
8. Ord K., Fildes R., Kourentzes N. Principles of Business Forecasting. Wessex, Incorporated, 2017. 588 p.
9. Svetunkov, I., Kourentzes, N. Complex Exponential Smoothing. Working Paper of Department of Management Science, Lancaster University 2015, 31 p.

10. Svetunkov Sergey. Complex-Valued Modeling in Economics and Finance. Springer Science+Business Media, New York, 2012. 318 p.
11. Vu Ky M. The ARIMA and VARIMA Time Series: Their Modelings, Analyses and Applications. AuLac Technologies Inc., 2007. 488 p.