Семенычев В.К. Новые возможности параметрического инструментария эконометрики на относительно коротких выборках. // Управление экономикой: методы, модели, технологии: четырнадцатая международная конференция: сборник научных трудов. Том 1 / Уфимск. гос. авиац. техн ун-т. – Уфа: УГАТУ, 2014, Т.1. – С.14-17.

В.К.Семенычев

НОВЫЕ ВОЗМОЖНОСТИ ПАРАМЕТРИЧЕСКОГО ИНСТРУМЕНТАРИЯ ЭКОНОМЕТРИКИ НА ОТНОСИТЕЛЬНО КОРОТКИХ ВЫБОРКАХ

МБОУ ВО «Самарская академия государственного и муниципального управления»

Рассмотрены возможности комплекса из десятков существенно нелинейных, многопараметрических и зачастую многокомпонентных моделей временных рядов показателей социально-экономических систем и методы их идентификации на основе обобщенных параметрических моделей авторегрессии-скользящего среднего, генетического алгоритма, Левенберга-Марквардта и других численных методов, отличительной особенностью которых является использование относительно коротких (до 50) наблюдений.

Ключевые слова:Нелинейные, многопараметрические и многокомпонентные параметрические модели, временные ряды, относительно короткие выборки, идентификация, моделирование, прогнозирование

Актуальностью обладает в настоящее время решение задачи моделирования и прогнозирования многопараметрических (содержащих более трех параметров), существенно нелинейных и зачастую многокомпонентных временных рядов, содержащих мультитрендовые и колебательные (циклические и сезонные) компоненты.

Особенно сложна задача обеспечения удовлетворительной точности решения этой задачи при эволюции анализируемых трендов, а также амплитуд колебательных компонент на относительно коротких выборках: от 10-12 до 50 наблюдений, причем при их комплексном (аддитивно-мультипликативном) взаимодействии с трендами.

Именно такая задача часто встречается в эконометрической практике при анализе «молодых», инновационных процессов, эволюционной динамики жизненного цикла продукта (ЖЦП): товара, товарной группы, услуги, бренда, организации и т.п.

К их числу относится и определение характеристических показателей динамики развития информационных технологий, добычи и потребления ЖЦП невосполняемых ресурсов (например, нефти, газа, угля, золота): пика, точки перегиба, определяющей нерасширяемый спрос, и асимптот.

Динамика таких ЖЦП зачастую имеет существенное экономическое значение, определяющее бюджеты стран.

Казалось бы, непараметрические (алгоритмические) модели обладают высокой универсальностью, но предложенный комплекс из многих десятков аналитических (параметрических) моделей, не уступая им в универсальности, позволил решить и новые задачи по долее «тонкому» и точному моделированию, прогнозированию при точечной и интервальной оценке точности параметров предлагаемых моделей, самих траекторий показателей, а также по определению моментов смены моделей, т.е. по мониторингу их эволюции.

Моделирование ЖЦП, а зависимости от конкретных приложений, может проводиться десятками логистических моделей: кумулятивных (накопленных) значений анализируемых показателей, получаемых из решения соответствующих дифференциальных уравнений динамики социально-экономических систем, или из алгебры взаимодействия кумулятивных моделей.

Кроме того,анализ эволюционной динамики может осуществляться импульсными моделями, например, Хабберта, Капицы-Коши, Маккея, логнормальной и др., а также феноменологическими (эмпирическими) кумулятивными и импульсными моделями, алгебраическим взаимодействием моделей всех указанных видов, включением кумулятивных моделей в импульсные, для приближения последних к реальной асимметрии на практике [1,2,3].

Среди используемых моделей параметрического комплекса есть известные, есть обобщения известных, а также - новые.

В отличии от немногочисленных известных решений подобных задач, достигаемых при некотором определенном (может быть даже единственном) наборе параметров моделей, поставлена и решена более общая постановка: определена точность применяемых методов идентификации в достаточно широком динамическом диапазоне всех параметров моделей и, более того, при дисперсии стохастической компоненты, неизбежно присутствующих в наблюдениях временного ряда, в диапазоне 0 - 30% от дисперсии модельных анализируемых значений ряда.

Делается это расчетом дисперсии модельных значений ряда, генерируемых программно, а также генерированием и стохастической компоненты, которую затем центрируют, рассчитывают ее среднеквадратического отклонения, делят на его значение (т.е. нормируют).

Полученные значения стохастической компоненты с нулевым математическим ожиданием и единичной дисперсией умножают на коэффициент сигнал-шум для задания соотношения дисперсий полезного сигнала (модели) и шума (стохастической компоненты) в диапазоне значений 0 - 30%, который традиционно принят в измерительной технике для принятия решения о практической работоспособности того или иного метода идентификации.

Принятая реализация численного эксперимента обусловлена тем, что известные программы реализации стохастической компоненты не позволяют гарантировать (назначать) известную дисперсию и обеспечивать симметричность ее распределения на коротких (до 50 наблюдений) выборках

Количество генерируемых тестовых реализаций в предлагаемой методикеназначают из условия вычислительной устойчивости идентификации и составляет, как правило, десятки тысяч для каждой из анализируемых моделей при каждом наборе параметров и соотношений дисперсий полезный сигнал-шум.

Таким образом, предложенной методикой анализа точности удается оценить область применения используемых методов идентификации моделей из условия обеспечения ими удовлетворительной точности моделирования, оцениваемой коэффициентом детерминации, в диапазоне значений от 70 до 99,9%.

Приемлемую точность прогнозирования определяли по значениям MAPE-оценки или второго коэффициента Тейла в диапазоне до 5-10%

Сформулированы также требования к минимальному объему используемой выборки для достижения указанных критериев точности, и по ее расположению на падающих, и растущих ветвях логистических кумулятивных и импульсных моделей ЖЦП (выборка должна, как правило, включать в себя точку перегиба траектории, а точностные характеристики идентификации на ветвях различаются).

Общее число моделируемых и прогнозируемых нами аналитических моделей ЖЦП превышает две сотни по количеству трендов и, кроме того, оно может быть существенно расширено за счет моделей колебательных циклических и сезонных компонент, в которых допускается эволюция их амплитуд по линейному, экспоненциальному и по другим законам.

Значительное внимание уделено анализу кумулятивных логистических моделей не только с фиксированной асимметрией, но и с произвольной асимметрией.

Кумулятивные логистические модели включались в импульсные кумулятивные модели для управления их асимметрией, обычно характерной для многих приложений.

Выбор той или иной модели временного ряда делался по лучшим значениям критерия точности моделирования и прогнозирования или комплексно: по их разности.

В зависимости от конкретного приложения задачи (важности моделирования или прогнозирования) разность может назначаться с весовыми коэффициентами, которые в сумме дают единицу.

При наличие колебательной компоненты в структуре временного ряда большую точность прогнозирования дает именно ее учет, определение ее структуры взаимодействия с трендом.

Рассмотрены известные и предложены новые структуры взаимодействия колебательных компонент с трендами.

Наиболее часто и просто считать колебательную компоненту ряда динамикине взаимодействующей с трендом (независимой от него) и входящей аддитивно в структуру ряда динамики

$$Y_k = T_k + S_k^A + \varepsilon_k ,$$

где Y_k - уровни моделируемого показателя ЖЦП, T_k - уровни тренда, S_k^A - уровни аддитивной колебательной компоненты, ε_k - стохастическая компонента, $k = 1, 2, ..., n$ - номера наблюдений ряда динамики, n - объем выборки.

Известная классическая структура ряда динамики с мультипликативной колебательной компонентой S_k^A

$$Y_k = T_k \cdot S_k^M \cdot \varepsilon_k ,$$

которая весьма затруднена для идентификации в силу того, что она оперирует долями (пропорциями) колебательной и стохастической компонент.

Показано, что большое практическое применение, особенно для макроэкономических показателейЖЦП, может найти параметрическая структура пропорционально-мультипликативно взаимодействия колебательной компоненты с трендом

$$Y_k = T_k \left(1 + S_k^M\right) + \varepsilon_k = T_k + T_k S_k^M + \varepsilon_k ,$$

где $\left|S_k^M\right| \le 1$.

Оправдана для многих случаев и «взвешенная по амплитуде аддитивно-мультипликативная модель» взаимодействия:

$$Y_k = T_k + S_k^{AM} + \varepsilon_k ,$$

где

$$S_k^{AM} = \sum_{i=1}^{N} A_i \left(\left[1 - \gamma_i\right] + \gamma_i \frac{T_k}{\max\left(T_k\right)} \right) \sin\left(\omega_i k \Delta + \varphi_i\right).$$

Веса γ_i в данной модели «привязаны» к частотам ω_i колебательной компоненты, а параметр $\max\left(T_k\right)$ обеспечивает нормировку для о возможности сравнения амплитуд аддитивной и мультипликативной колебательных компонент в одном диапазоне значений.

Эту модель можно рассматривать и как обобщение аддитивной структуры (при $\gamma_i = 0$) или пропорционально-мультипликативной структуры (при $\gamma_i = 1$).

Ее свойства демонстрирует представление в виде суммы гармоник

$$S_k^{AM} = \sum_{i=1}^{N} S_k^{AM\,i} ,$$

где

$$S_k^{AM\ i} = A_i \left(\left[1 - \gamma_i\right] + \gamma_i \frac{T_k}{\max\left(T_k\right)} \right) \sin\left(\omega_i k \Delta + \varphi_i\right)$$

.

Каждая гармоника в этом случае будет суммой пропорционально-мультипликативной

$$S_k^M = A_i \frac{T_k}{\max\left(T_k\right)} \sin\left(\omega_i k\Delta + \varphi_i\right)$$

и аддитивной

$$S_k^A = A_i \sin\left(\omega_i k\Delta + \varphi_i\right)$$

гармоник с частотой ω_i, взятой, соответственно, с весами γ_i и $1-\gamma_i$.

В отдельных случаях колебательная компонента ЖЦП может менять свою частоту в зависимости от величины уровня тренда.

По аналогии с предыдущим результатом такую структуру можно назвать «взвешенной по частоте колебательной компонентой»:

$$S^{\Omega_1} = \sum_{i=1}^{N_\Omega} A_i \left[1+\gamma_i\left(\frac{T_k}{T_{\max}}-1\right)\right] \sin\left(\omega_i \int_{t_0}^{t_k} \left[\frac{T(s)}{T_{\max}}\right]^{-\theta_i} ds + \varphi_i\right),$$

$$S^{\Omega_2} = \sum_{i=1}^{N_\Omega} A_i \left(\frac{T_k}{T_{\max}}\right)^{\gamma_i} \sin\left(\omega_i \int_{t_0}^{t_k} \left[\frac{T(s)}{T_{\max}}\right]^{-\theta_i} ds + \varphi_i\right).$$

Частота и амплитуда в такой модели могут быть постоянными, а могут и зависеть от тренда. Последняя модель отражает возможность изменения частоты колебательной компоненты при малых уровнях тренда.

Напомним, для аналогии, что ив физике тела с большей массой устойчивее при гармонических воздействиях.

Обоснование и многие количественные примеры приведенных структур взаимодействия трендов ЖЦП с колебательными компонентами для анализа динамики добычи невосполняемых ресурсов (нефти, газа, угля, золота) приведены в [4].

Идентификация комплекса моделей и решение задачи определения возможной области их приложения потребовало также применение и развитие аналитических и численных методов.

Исторически первым из разработанных аналитических методовидентификации оказался метод обобщенных параметрических моделей авторегрессии-скользящего [5], особенностью которого является конструирование моделей авторегрессии-скользящего среднего при помощи Z-преобразования (преобразования Лорана) детерминированных компонент моделей ряда с коэффициентами, определяемыми через параметры анализируемых моделей.

Процедура идентификации является двухэтапной. На каждом этапе применяется метод наименьших квадратов (МНК) для расчета коэффициентов авторегрессий и связанных (нелинейно) с ними параметров моделей.

При реализации данного метода необходимо на первом этапе идентификации использовать приемы снижения автокорреляций наблюдений: для обеспечения условий Гаусса-Маркова и получения оптимальных оценок параметров моделей.

В ряде случаев оказалось необходимо компенсировать гетероскедастичность стохастической компоненты.

Оправдано конструирование обобщенных параметрических моделей авторегрессии-скользящего среднего с помощью программы MAPLE.

Вычислительная устойчивость данного метода ограничена 4-6 порядками конструируемых обобщенных параметрических моделей авторегрессии –скользящего среднего.

Для некоторых моделейряда динамики нормальная система алгебраических уравнений при применении МНК оказывалась нелинейной, что заставило искать аналитическое решение с использованием базиса Гребнера [6]. Метод базисов Гребнера сводит систему нормальных уравнений к треугольному виду, а задачу ее решения – к решению последовательности алгебраических уравнений, что легко сделать программным путем.

Использование Z- преобразования для конструирования обобщенных параметрических моделей авторегрессии-скользящего среднего возможно далеко не для всех моделей ЖЦП и рядов динамики, поэтому были рассмотрены и применения численных методов.

Оказалось, что лучшими характеристиками во многих случаях обладают методы Левенберга-Марквардта и метод генетического моделирования (ГА), который оказался наиболее универсальным по видам рассматриваемых моделей и по структуре вхождения стохастической компоненты.

ГА используется для решения задач оптимизации и моделирования путём случайного подбора, комбинирования и вариации искомых параметров с использованием механизмов, напоминающих биологическую эволюцию.

Его сходимость и набор найденных корней не зависят от выбора начального приближения или диапазона допустимых значений корней. Для идентификации был использованы настройки генетического алгоритма, реализованного функциями gatool системы MATLAB.

Для тренд-колебательных рядов динамики высокие результаты по точности и универсальности показал предложенный метод «итерационной тренд-сезонной декомпозиции» [7].

Его достоинствами, по сравнению с известной «классической декомпозицией», является использование параметрических моделей на поочередных итерационных этапах «десезонализации» и «детрендирования», возможность изменения их очередности в зависимости от преобладания дисперсии этих компонент, а также при эволюции амплитуд колебательных компонент и довольно быстрая сходимость.

Остановимся еще на одном достоинстве предложенного инструментария.

При решении задачи определения области применения комплекса используемых моделей и соответствующего им комплекса методов идентификации с использованием тестовых моделей и стохастических компонент возможно получение точечных и интервальных оценок точности всех оценок параметров, точечных и интервальных оценок коэффициентов детерминации и точности прогнозирования.

Это означает возможность задания доверительной вероятности и расчета доверительного интервала для каждой из точек траектории ряда.

Трансформацию идеи тестовых выборок можно использовать и для определения доверительного интервала имеющейся выборки наблюдений до наблюдения Y_k. Для этого по имеющейся выборке находят невязки для всех наблюдений от первого до *к*-го. Затем рассчитываем их дисперсию, генерируем программно (центрированные и нормированные) выборки с такой же дисперсией и находим доверительный интервал для Y_k.Потом задаем горизонт прогноза и определяем до каких последующих наблюдений ряда динамики ($k+1, k+2, ..., n$) прогнозные значения оказываются внутри этого интервала.

Тем самым возможно осуществлять мониторинг эволюции модели (по параметрам или по виду), фиксировать момент смены модели при выходе прогнозного значения за границы доверительного интервала.

При этом оказывается возможной осуществление мультимодельного («склеенного» из нескольких видов) моделирования, что существенно может увеличить точность прогнозирования временных рядов [8].

Важным представляется тообстоятельство, что предложенные комплексы моделей и методов идентификации реализованы для многих приложений (добыча и потребление натуральных показателей невосполняемых ресурсов, спрос на товары и товарные группы, спрос на продукт, доли рынка, динамика безработицы и эмиграции и др.), а также для различных уровней агрегирования анализируемых показателей (для предприятий, регионов, стран, группы стран).

Получены решения по анализу и мультитрендовых структур рядов динамики, включающих в себя несколько взаимосвязанных или независимых трендов: например, аддитивных (в том числе повторных ЖЦП), мультипликативных или пропорционально-мультипликативных.

Такие структуры позволяют описывать еще более сложную динамику, когда-либо один тренд постепенносменяется другим, либо формируется новая динамика, существенно отличающаяся от динамики исходных трендов.

При мультитрендовых моделях возможны и новые взаимодействия с колебательными компонентами, как сезонными, так и циклическими.

Разработаны и защищены авторскими свидетельствами комплексы программных средств, внутри которого «скрыта» довольно сложная математика, а интерфейс является дружественным по отношению к пользователю, обладает многими регулируемыми настройками [10, 11 идр.].

Библиографический список

1.Семенычев В.К., Семенычев Е.В. Параметрическая идентификация рядов динамики: структуры, модели, эволюция. Самара. Изд-во «Сам НЦ РАН», 2011. 364 с.

2.Семенычев Е.В. Эконометрическое моделирование жизненного цикла продукта. - Самара, САГМУ, 2012. 148 с.

3.Семенычев В.К., Кожухова В.Н. Анализ и предложения моделей экономической динамики с кумулятивным логистическим трендом. Самара. Изд-во «Сам НЦ РАН», 2013. 152 с.

4.Семенычев В.К., Куркин Е.И., Семенычев Е.В., Данилова А.А Инструментарий моделирования колебательной компоненты в колоколообразных кривых жизненного цикла продукта. Прикладная эконометрика. 2014. №33 (1), С. 111 - 124.

5.Семёнычев В.К. Идентификация экономической динамики на основе моделей авторегрессии. Самара: Изд-во СНЦ РАН, 2004. – 243 с.

6. Семёнычев В.К., Семёнычев Е.В., Коробецкая А.А. Метод параметрической итерационной декомпозиции тренд-сезонных рядов аддитивной структуры // Вестник Самарского муниципального института управления. 2010. №1(12). С. 63-72.

7.Семенычев В.К., Куркин Е.И., Семенычев Е.В. Идентификация жизненного цикла продукции на основе моделей авторегрессии-скользящего среднего и базисов Гребнера. Прикладная эконометрика. №1(25) .2012. – С.122-137.

8.Семенычев Е.В., Куркин Е.И., Семенычев В.К. Мультимодельная оценка эволюции жизненного цикла при добыче невозобновляемых ресурсов. // Проблемы экономики и управления нефтегазовым комплексом. М., «ВНИИОЭНГ», 2014.№10 – С.28-34.

9.Семенычев В.К., Коробецкая А.А. Идентификация мультитрендовой эволюционирующей модели с мультипликативной линейно-экспоненциальной структурой. Вестник Самарского муниципального института управления. Изд-во САГМУ. Самара. 2014. №1(28), - С. 7 - 14.

10.Свидетельство о регистрации программы для ЭВМ «EconometricResearch» №2008610493 от 25 января 2008 г./ Семенычев В.К., Семенычев Е.В, Сергеев А.В., Маркина О.С.

11.Свидетельство о государственной регистрации программы для ЭВМ «Программа «ZGroebnerident» №2011615047 от 29 июня 2011 г. / Семенычев В.К., Семенычев Е.В., Куркин Е.И.