О.В. Тимченко

Технологии проведения предпрогнозного анализа социально-экономической сферы

На сегодняшний момент база инструментальных и математических методов прогнозирования дискретных эволюционных процессов достаточно разнообразна. Наибольший интерес, на наш взгляд, представляют модели, которые сочетают в себе технологии клеточных автоматов, фрактального анализа и теории нечетких систем.

Инструментальную базу предлагаемой прогнозной модели составляют клеточные автоматы, память которых отражает долгосрочную память рассматриваемого временного ряда. Фрактальный анализ предшествует построение модели прогноза рассматриваемого временного ряда и позволяет выявить наличие в ряде долгосрочной памяти. Прогноз в этом случае можно получить не только в традиционном числовом выражении, но и с использованием терминологии теории нечетких множеств.

Предпрогнозный анализ играет огромную роль при определении параметров долгосрочной памяти временного ряда. Целью фрактального анализа как инструмента предпрогнозного анализа являются оценка глубины, трендоустойчивости, цикличности, выявление показателя Херста долгосрочной памяти временного ряда [1: 269]. Знание перечисленных фрактальных характеристик рассматриваемого временного ряда представляет аналитику предпрогнозную информацию, т.е. позволяет ему оценить перспективность надежного прогнозирования временного ряда с помощью клеточноавтоматной прогнозной модели.

Основным инструментарием фрактального анализа временного ряда является алгоритм R/S – анализа. Суть этого анализа в следующем.

Рассмотрим выборку из $n+1$ элементов (здесь n – четно). Тогда выборку можно рассматривать как временной ряд, причем число уровней ряда достаточно велико (например, Эдгар Петерс при исследовании фондового рынка брал около 5 000 значений) [4:254]. Подвергнем выборку дальнейшим преобразованиям с помощью логарифмических соотношений, что позволит получить временной ряд со значениями уровней длиной n: $\log\frac{U_i}{U_{i-1}}$.

Далее для n найдем наименьший собственный делитель m, не меньший 10. Разделим логарифмический временной ряд на $k=\frac{n}{m}$ групп. Элементы в каждой группе обозначим t_i. Для каждой группы находим средние значения

$$\overline{t_k}=\frac{1}{m}\sum_{i=1}^{m}t_i,\quad k=1...m$$

$$\overline{t_k}=\frac{1}{m}\sum_{i=m+1}^{2m}t_i,\quad k=m+1...2m$$

$$\overline{t_k}=\frac{1}{m}\sum_{i=(k-1)m+1}^{n}t_i\quad,\quad k=(k-1)m+1...n$$

и накопленные отклонения от среднего значения

$$X_1=t_1-\frac{1}{m}\sum_{i=1}^{m}t_i,$$

$$X_2=(t_2-\frac{1}{m}\sum_{i=1}^{m}t_i)+X_1\quad ...$$

$$X_m=(t_m-\frac{1}{m}\sum_{i=1}^{m}t_i)+X_{m-1}$$

Затем по каждой группе находим нормированный размах

$$R_i=\max(X_i)-\min(X_i)$$

и стандартное отклонение

$$S_i=\left(\frac{1}{m}\sum_{i=1}^{m}(t_i-\overline{t_k})^2\right)^{1/2}$$

Коэффициент R/S для каждой группы рассчитывается как R_k / S_k. Затем находим средний размах вариации

$$\overline{\left(\frac{R}{S_j}\right)} = \frac{1}{k}\sum_{i=1}^{k}\frac{R}{S_i}$$

Смысл индекса j состоит в том, что на j шаге получен средний размах вариации, который соответствует собственному делителю такого же порядка. Описанная процедура повторяется до тех пор, пока не будут выбраны все собственные делители. На последнем шаге $m = \frac{n}{2}$.

В результате будет получена выборка $\overline{\left(\frac{R}{S_j}\right)}, j = 1 \ldots \frac{n}{2}$, для которой количество содержащихся элементов равно количеству собственных делителей n. На основе построенной выборки можно составить уравнение линейной регрессии в логарифмической системе координат, т.е. факторный признак совпадает с логарифмом количества элемент в группе j

$$\log\frac{R}{S} = \log c + H \cdot \log k,$$

а зависимой переменной является логарифм показателя $\frac{R}{S}$. Здесь $\acute{I}$ – показатель Херста, $\tilde{n}$ – мера корреляции.

Такое уравнение является линеаризованным, и его параметры удобно найти методом наименьших квадратов. Проверка значимости уравнения регрессии проводится традиционным образом на основании F-критерия Фишера. Значимость параметра $\acute{I}$ может быть проверена с помощью t-статистики Стьюдента [2:21].

Показатель Херста имеет важное значение для определения фрактальных характеристик временного ряда – цвета шума, который будет соответствовать данному временному ряду на том или ином отрезке времени.

В зависимости от значений показателя Херста принято различать шум трех цветов:

- если $0 < \acute{I} < 0{,}5$, то для ряда характерен «розовый» шум. В этом случае говорят о антиперсистентном ряде, т.е. ряде для которого более вероятна смена предыдущего направления. Такие процессы, как правило, сопровождаются эффектами турбулентности;
- если $\acute{I} = 0{,}5$, то для ряда характерен «белый» шум. В этом случае говорят о стохастичном ряде;
- если $0{,}5 < \acute{I} < 1$, то для ряда характерен «черный» шум. В этом случае говорят о персистентном ряде, т.е. ряде, для которого более вероят-

на трендовость (сохранение направленности). Такие ряды, как правило, и наблюдаются на финансовых рынках и в социально-экономических процессах.

Кроме того, можно выделить нечеткую границу между «белым» и «черным» шумом, так называемый «серый» шум.

Изучение $\acute{I}$ -траектории позволяет сделать вывод о наличие долговременной памяти рассматриваемого временного ряда, если $\acute{I}$ -траектория через несколько своих начальных точек оказывается в области черного шума, а его R/S – траектория в этих же точках имеет трендовую направленность, тогда глубину долгосрочной памяти определяет такой номер l, для которого соответствующей точке траектория показателя $\acute{I}$ получает отрицательное приращение, а R/S-траектория в этой точке демонстрирует так называемый «срыв с тренда», т.е. резкое изменение тренда [3: 11].

Если в данном временном ряде случайным образом перетасовать его элементы и полученный ряд представить на вход алгоритма R/S-анализа, то на выходе алгоритма максимальное значение показателя Херста и R/S-траектории окажется значительно меньше, по сравнению со значениями для исходного временного ряда, в случае если этот временной ряд обладает долговременной памятью.

Понятия «глубина памяти» и «длина квазицикла» тесно связаны в смысловом отношении, т.е. по своей содержательной сути одно из этих понятий обуславливает другое и наоборот. Вместе с тем, эти термины представляют собой хотя и родственные (взаимно обусловленные), но различные понятия.

Термин «квазицикл» близок к термину «цикл» в общепринятом понимании. Различие между этими терминами состоит в том, что начальная и конечная точки квазицикла не обязаны совпадать. Вхождение некоторой точки в сравнительно малую окрестность начальной точки и определяет окончание квазицикла. При этом может наблюдаться самопересечние начального и конечного звеньев квазицикла, в том случае если это приводит к наилучшему сближению граничных точек.

Как показывают экспериментальные расчеты, в процессе проведения фрактального анализа временных рядов особого внимания заслуживают те из них, в которых длина имеющихся квазициклов сравнима с длиной самого ряда.

Выявленный с помощью фрактального анализа факт наличия памяти в рассматриваемых социально-экономических временных рядах является основанием для того, чтобы в целях прогнозирования временного ряда использовать такие математические модели, которые формируют результат прогнозирования на базе учета специфики выявленной памяти. В процессе использования методов нелинейной динамики (фрактальный анализ, фазовый анализ) получаем предпрогнозные результаты в виде количественных и качественных оценок свойств динамики рассматриваемых временных рядов [5: 212]. Эти свойства по своей сути являются системными, т.е. они не присущи отдельным элементам системы и в то же время представляют информацию о тенденциях и закономерностях в поведении исследуемой системы. К числу таких системных свойств можно отнести полученную в процессе предпрогнозного анализа информацию о наличии долговременной памяти, оценки ее глубины, а также оценки значений показателя Херста и степень близости его к случаям хаотического поведения.

Приведенная схема предпрогнозного анализа социально-экономических процессов применима к реальным процессам. Но эффективность анализа значительно возрастет, если в рамках этой схемы использовать дополнительные модели фазового анализа, а саму модель строить с помощью клеточного автомата.

Библиографический список

1. Воробьев Г.А., Павленко И.И. Тенденции и перспективы информатизации социального управления в современных условиях // Гуманитарий Юга России. 2016. Т. 20. № 4. С. 144-151.
2. Давнис В.В., Тимченко А.Б. Прогнозирование волатильности финансовых показателей: адаптивный подход // Вестник Воронежского государственного университета. Серия: Экономика и управление. 2006. № 2. С. 269-277.
3. Зиненко А.В. R/S анализ на фондовом рынке // Бизнес-информатика. 2012. № 3. С. 21.
4. Кратович П. В. Предпрогнозный анализ временных рядов финансовых данных на основе методов фрактального анализа // Молодой ученый. 2010. № 1-2. Т. 1. С. 11-18. URL: https://moluch.ru/archive/13/1084/ (дата обращения: 23.01.2018).
5. Мандельброт Б., Хадсон Р. (Не)послушные рынки. Фрактальная революция в финансах. М.: Изд. дом «Вильямс», 2006.
6. Суюнова Г.Б., Гайворонская Н.А., Тимченко О.В., Тимченко А.Б. Моделирование прогнозных оценок стоимости базовых активов с помощью метода авторегрессии // Научное обозрение. 2015. № 18. С. 212-218.
7. Vorobyev G.A., Ryndjuk V.A., Kozlov V.A., Makarov A.M. Probabilistic models of cryptographic systems and their applications. В сборнике: 2016 3rd International Conference on Digital Information Processing, Data Mining, and Wireless Communications, DIPDMWC 2016 3. 2016. С. 160-163.