Методы определения почти – периодов в эмпирических данных с трендом

В.И. Кузьмин - доктор технических наук, профессор, кафедра «Прикладная математика».
А.Ф. Гадзаов - аспирант кафедры «Прикладная математика»

Построение моделей по эмпирическим данным часто связано трудностями выделения колебаний, сопутствующих трендам. Ньютон предложил решать задачу о разделении движений для механических систем [1], где тренд определяется уравнениями движения центра масс и отклонения от него реальной траектории позволяют отдельно анализировать колебания. Именно этим методом воспользовался Н.Д.Кондратьев при анализе данных по экономической динамике, что привело его к открытию больших циклов конъюнктуры, называемых сегодня «волнами Кондратьева» [2].

На рис. 1 представлены данные о долге вкладчикам сберегательных касс во Франции, которые анализировал Н.Д. Кондратьев, аппроксимированные параболой методом наименьших квадратов. На оставшихся после исключения тренда колебаниях, представленных на рис.2, можно отметить наличие периода длительностью пятьдесят семь лет.

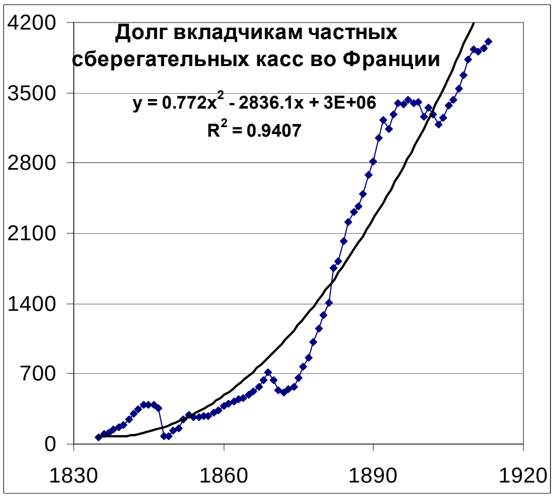

Рис.1 Долг вкладчикам частных сберегательных касс во Франции

Рис.2 Данные рис.1 после исключения тренда

В случае, когда уравнения трендов не известны, определение вида трендовой зависимости представляет большие сложности, что приводит, в ряде случаев, к неопределенности получаемых для колебаний результатов.

Рассмотрим возможности решения этой задачи. За основу возьмем линейное неоднородное дифференциальное уравнение первого порядка с переменными коэффициентами (1),

$$\dot{y}(x) + p(x)\cdot y(x) = q(x) \quad (1)$$

Учитывая что, при известных частных решениях уравнения (1) $y_1(x), y_2(x)$

$$y(x) = y_1(x) + C(y_2(x) - y_1(x)) \ (2),$$

или частное решение $y_3(x)$ определяется соотношением:

$$y_3(x) = y_1(x) + C(y_2(x) - y_1(x)) \ (3)$$

Таким образом, решения уравнения (1) обладают следующими геометрическими свойствами [3], при которых всякая интегральная кривая линейного уравнения (1) делит в постоянном отношении отрезок ординаты между какими-либо двумя интегральными кривыми этого уравнения (4), рис. 3.

$$\frac{N_2N_3}{N_3N_1} = \frac{M_2M_3}{M_3M_1} \ (4)$$

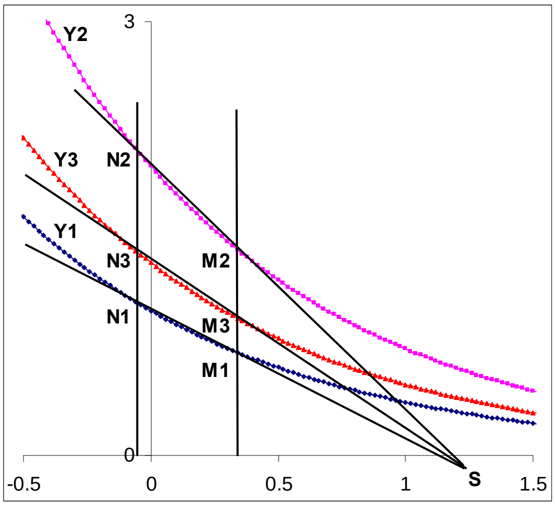

Рис.3 Иллюстрация геометрических свойств решений (1)

Перепишем уравнение (3), как

$$\frac{y_3(x) - y_2(x)}{y_2(x) - y_1(x)} = (C_1 - 1) = K \ (4)$$

Так как (3) в общем виде можно записать:

$$\frac{y_{общ.одн}(x) + y_3(x) - y_{общ.одн}(x) - y_2(x)}{y_{общ.одн}(x) + y_2(x) - y_{общ.одн}(x) - y_1(x)} = (C_1 - 1) = K \ (5)$$

Где, $y_{общ.одн}(x)$ - решение общего неоднородного уравнения.

Вместо частных решений $y_1(x), y_2(x), y_3(x)$ возьмем $y_{исх}(x-\tau), y_{исх}(x), y_{исх}(x+\tau)$ исходного эмпирического ряда и перепишем уравнение (4)

$$\frac{y_{исх}(x+\tau) - y_{исх}(x)}{y_{исх}(x) - y_{исх}(x-\tau)} = (C_1 - 1) = K \ (6)$$

Таким образом, при К = const, уравнение (6) исключит из эмпирических данных трендовую компоненту, что позволит выделить колебания для дальнейшего анализа их структуры.

Использование соотношения (6) бывает не всегда удобно, поэтому преобразуем его к виду:

$$y_{исх}(x+\tau)-y_{исх}(x)=K\cdot(y_{исх}(x)-y_{исх}(x-\tau)) \quad (7)$$

Возьмем К = 1, так как при таком значении обеспечивается постоянство вторых разностей уравнения (7). Тогда

$$\frac{y_{исх}(x+\tau)-y_{исх}(x-\tau)}{2\cdot y_{исх}(x)}=1 \text{ или } \ln\left(\frac{y_{исх}(x+\tau)-y_{исх}(x-\tau)}{2\cdot y_{исх}(x)}\right)=0 \quad (8)$$

Взяв, вместо $y_{исх}(x)$ $\ln(y_{исх}(x))$ получим условие постоянства относительных приростов. Тогда уравнение (7) преобразуется:

$$\ln\left(\frac{y_{исх}(x+\tau)\cdot y_{исх}(x-\tau)}{y_{исх}^2(x)}\right)=0 \quad (9)$$

Таким образом в следующих координатах исключается тренд

$$\ln\left(\frac{y_{исх}(x+\tau)-y_{исх}(x-\tau)}{2\cdot y_{исх}(x)}\right)-t \quad (10),$$

$$\ln\left(\frac{y_{исх}(x+\tau)\cdot y_{исх}(x-\tau)}{y_{исх}^2(x)}\right)-t \quad (11)$$

Задачей остается определение таких значений τ, которые соответствуют параметрам кусочно-экспоненциальных участков тренда. Определение значений τ может быть реализовано путем построения функций, характеризующих величину среднего значения колебаний, полученных после исключения тренда. Эффективными значениями τ будут считаться такие, которые дают нулевую величину среднего значения для колебаний.

Для выявления периодов, свободных по возможности от априорных предположений, остается фундаментальное свойство периода функции.

Пусть имеется строго периодическая функция $f(t)$ с периодом τ. Тогда

$$f(t+\tau)-f(t)=0. \quad (12)$$

Если функция не строго периодическая, то разность $f(t+\tau)-f(t)$ для разных τ будет принимать различные значения. Введем среднее значение абсолютной разности для каждого возможного периода τ. Для дискретного случая, если n общее число отсчетов функции $f(t)$, заданной экспериментальными значениями, то средняя абсолютная разность для некоторого пробного периода τ будет равна

$$a(\tau) = \frac{1}{n-\tau} \sum_{t=1}^{n-\tau} |f(t+\tau) - f(t)|. \qquad (13)$$

Эта функция называется сдвиговой или функцией Джонсона [4].

Значениями τ наиболее близкие к периоду определяются положением минимумов функции $a(\tau)$. Они называются почти – периодами.

Главный почти - период τ_1 функции $f(t)$ может быть определен как

$$\tau_1 = \arg\min_{\tau_{\min} \le \tau \le \tau_{\max}} a(\tau)$$

где $\tau_{\min}$ и $\tau_{\max}$ - естественные пределы поиска почти - периодов, выбираемые таким образом, что с одной стороны отбрасываются $\tau \prec \tau_{\min}$, при которых функция $a(\tau)$ может принимать малые значения из-за инерционности функции $f(t)$ и, с другой стороны, отбрасываются большие $\tau \succ \tau_{\max}$, при которых определение средней $a(\tau)$ становится ненадежным из-за малости числа $(n-\tau)$ - предела суммирования в выражении (13).

Фактически в этом случае мы получаем возможность выявлять почти - периоды, представленные в экспериментальных данных вне зависимости от формы колебаний.

Применение изложенных методов проиллюстрируем на данных ежемесячной динамике стоимости акций компании Дженерал – Моторс (General – Motors) [4] и сравним полученные результаты.

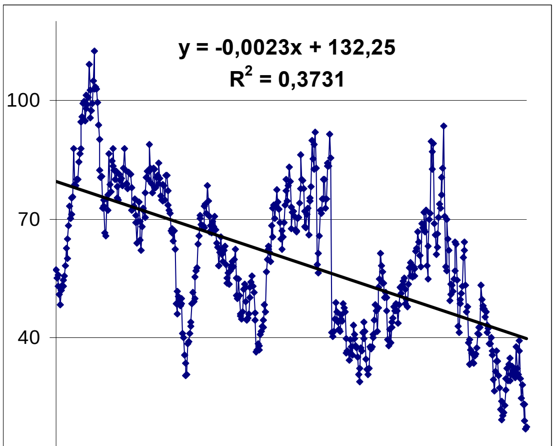

Рис. 4 Ежемесячные данные индекса Доу-Джонса, с 1900 г. по 2007 г., по оси ординат логарифмический масштаб, данные апроксимированы параболой методом наименьших квадратов

На рис.5а изображены данные после исключения тренда относительно параболы на рис.4. Сдвиговая функция, соответствующая этим колебаниям изображена на рис.5б. Варьирование величины τ для данных, представленных в координатах (10) и (11), показало, что среди эффективных значений есть 10 лет. Полученные после преобразования (10) и (11) результаты представлены на рис.6а и рис.7а. Сдвиговые функции, соответствующие этим колебаниям изображены на рис.6б и рис.7б.

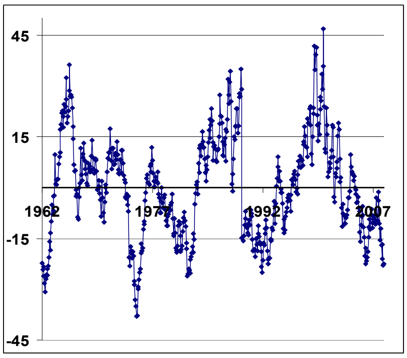

Рис.5а Данные после исключения тренда относительно параболы

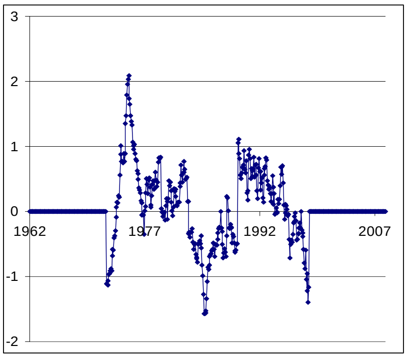

Рис.6а Данные после исключения тренда методом (10)

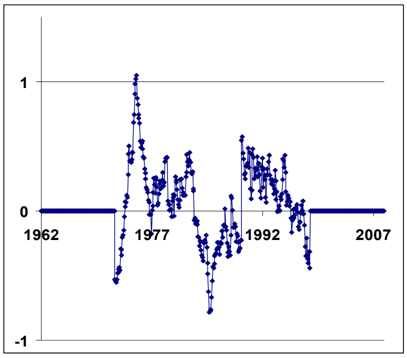

Рис.7а Данные после исключения тренда методом (11)

Рис.5б Сдвиговая функция для данных **рис.5а**

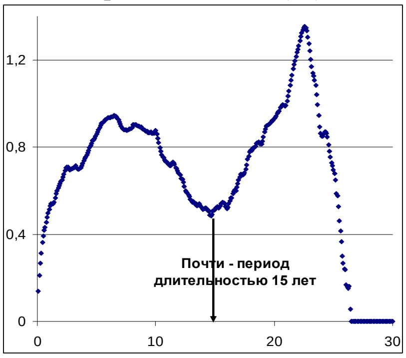

Рис.6б Сдвиговая функция для данных **рис.6а**

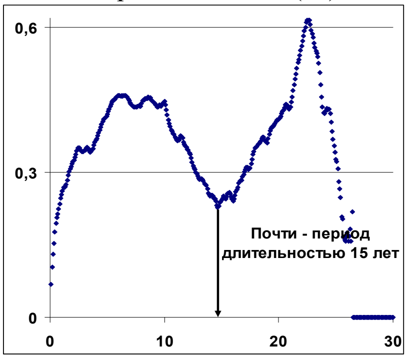

Рис.7б Сдвиговая функция для данных **рис7а**

Представленные на рис.5б, рис.6б, рис.7б результаты показывают наличие в динамике акций Дженерал – Моторс почти – периодов продолжительностью около 10 лет, который воспроизводится на всех перечисленных рисунках.

Рассмотренные алгоритмы, позволяет выделять почти – периоды в экспериментальных данных, с колебаниями относительно тренда.

Список литературы:

1. Ньютон И. Математические начала натуральной философии. - М.: Наука, 1989.
2. Кондратьев Н.Д. Проблемы экономической динамики. - М.: Экономика, Изд. ин. лит., 1989. Кондратьев Н.Д., Опарин Д.И. Большие циклы конъюнктуры. - М.: Институт экономики, 1928.
3. Матвеев Н. М. Методы интегрирования обыкновенных дифференциальных уравнений. – М.: Высшая школа, 1967.
4. Johnson M. Correlations of cycles in weather, solar activity, geomagnetic values and planetary configurations. - San Fransisco, Phillips and Van Orden, 1944
5. http://www.yahoo.com