УДК 519.2:303.732.4

01.00.00 Физико-математические науки

НЕПАРАМЕТРИЧЕСКИЕ ОЦЕНКИ ЦИКЛОВ

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор
РИНЦ SPIN-код: 4342-4994
Московский государственный технический университет им. Н.Э. Баумана, Россия, 105005, Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru

Во многих прикладных задачах рассматривают временной ряд, или случайный процесс), являющийся суммой детерминированной периодической функция от времени и случайных погрешностей, искажающих периодический сигнал. Требуется оценить длину периода и периодическую составляющую. При этом не предполагаем, что периодическая функция входит в какое-либо параметрическое семейство функций, например, конечных сумм синусов и косинусов. Очевидно, что предположение о вхождении периодической функции в параметрическое семейство не соответствует свойствам реального мира, т.е. является условным, внутриматематическим (ищем ключи под фонарем, потому что там светло, а не в кустах, где потеряли, потому что там темно). По аналогичным причинам нельзя предполагать, что функция распределения случайных погрешностей входит в какое-либо параметрическое семейство распределений. В соответствии с новой парадигмой математической статистики в настоящей статье рассматриваем задачу непараметрического оценивания (минимальной) длины периода и периодической составляющей сигнала. На основе естественных показателей разброса и размаха предлагаем новый класс непараметрических оценок длины периода и периодической составляющей во временных рядах. Исходя из общих результатов статистики объектов нечисловой природы доказана состоятельность этих оценок. С прикладной точки зрения необходимо численно минимизировать (по одному параметру -возможной длине периода) один или несколько из 66 описанных в статье функционалов

Ключевые слова: МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, ПРИКЛАДНАЯ СТАТИСТИКА, НЕПАРАМЕТРИЧЕСКАЯ СТАТИСТИКА, ВРЕМЕННОЙ РЯД, ДЛИНА ПЕРИОДА, ПЕРИОДИЧЕСКАЯ СОСТАВЛЯЮЩАЯ, СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ, СОСТОЯТЕЛЬНОСТЬ

UDC 519.2:303.732.4

Physics and mathematical sciences

NON-PARAMETRIC CYCLES ESTIMATORS

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci., professor
Bauman Moscow State Technical University, Moscow, Russia

In many applications, we study the time series (or a random process), which is the sum of the periodic deterministic function of time and random errors that distort the periodic signal. It is required to estimate the length of the period and the periodic component. It does not assume that the periodic function is included in any parameter family of functions, such as finite sums of sines and cosines. It is obvious that the assumption of occurrence of a periodic function in parametric family does not meet the characteristics of the real world, ie, is conditional, internal mathematical (look for the keys under the lamp because there is a light, not in the bush, where lost, because there are dark). For similar reasons, it is impossible to assume that the distribution function of the random errors is included in any parameter family of distributions. In accordance with the new paradigm of mathematical statistics in this article we studied the problem of nonparametric estimation (minimum) length of the period and the periodic component of the signal. On the basis of natural variation and scope of indicators is suggested a new class of nonparametric estimators of the length of the period and the periodic component in the time series. Based on the general results of statistics of objects of non-numeric nature we proved the consistency of these estimates. From the practical point of view it is necessary to minimize the numerical (one parameter - ability length of period of time) one or more of the 66 functionals, described in the article

Keywords: MATHEMATICAL STATISTICS, APPLIED STATISTICS, THEORETICAL STATISTICS, NONPARAMETRIC STATISTICS, TIME SERIES, LENGTH OF THE PERIOD, PERIODIC COMPONENT, STATISICAL ESTIMATION, CONSISTENCY

1. Введение

Во многих прикладных задачах рассматривают временной ряд (или случайный процесс) $y(t) = x(t) + e(t)$, где $x(t)$ - детерминированная периодическая функция от времени t, т.е. $x(t) = x(t + T)$ при некотором T, где T - длина периода (минимальная из возможных), а $e(t)$ - "шумы", случайные погрешности, искажающие периодический сигнал. Требуется оценить длину периода T и периодическую составляющую $x(t)$. При этом не предполагаем, что функция $x(t)$ входит в какое-либо параметрическое семейство, например, конечных сумм синусов и косинусов. Очевидно, что подобное предположение не соответствует свойствам реального мира, т.е. является условным, внутриматематическим (ищем ключи под фонарем, потому что там светло, а не в кустах, где потеряли, потому что там темно). В соответствии с методологией работ [1, 2] в настоящей статье мы рассматриваем задачу непараметрического оценивания (минимальной) длины периода и периодической составляющей сигнала.

Приведем примеры прикладных постановок, в которых необходимо решать подобную задачу.

1. По акустическим сигналам необходимо установить тип двигателя. Предполагается, что двигатели различаются по длине периода и виду основного периодического сигнала. Процедура идентификации основана на оценивании длины периода и периодической составляющей регистрируемого сигнала. Например, акустик подводной лодки должен идентифицировать источник сигналов, в частности, определить тип судна и его национальную принадлежность, с целью принятия управленческих решений.

2. В предположении цикличности экономических процессов требуется статистическим данным установить длину цикла и на основе вида периодической составляющей построить прогноз, например, прогноз

урожайности, емкости рынка тех или иных товаров или экономической активности в целом. Экономическим циклам посвящено довольно много публикаций [3 - 6]. К сожалению, в них не уделяется адекватного внимания методам выделения периодов и оценивания периодической составляющей; более того, зачастую нет даже алгоритмов выделения начала и конца цикла. Другими словами, циклы выделяются неформальными методами.

3. Для среднесрочного прогнозирования развития социокультурной сферы (социально-политического "климата", живописи, музыки, архитектуры, поэзии и т.д.) необходимо выявить ее цикличность [7] с помощью объективных измерений на базе субъективных первичных данных (т.е. на базе оценок экспертов).

4. В исторических событиях, описываемых согласно распространенной в настоящее время т.н. скалигеровской хронологии, обнаруживают цикличность [8], что полностью объясняется новой статистической хронологией [9], построенной с помощью методов статистики объектов нечисловой природы, и одновременно служит ей подтверждением.

2. Описание непараметрического метода оценивания

Пусть рассматриваемые функции $y(t)$, $x(t)$, $e(t)$ определены на отрезке $[0; A]$. При фиксированном T рассмотрим "куски" сигнала $y(t)$ на последовательных отрезках длины T, т.е. на отрезках $[0; T]$, $[T; 2T]$, $[2T; 3T]$, ... Удобно ввести последовательность функций на отрезке $[0; T]$, полученную сдвигами этих кусков к началу координат: $y_1(t) = y(t)$, $y_2(t) = y(t + T)$, $y_3(t) = y(t + 2T)$, ... Все они определены на отрезке $[0; T]$. Число этих функций равно числу полных периодов длины T, укладывающихся на отрезке $[0; A]$, т.е. равно целой части числа A/T. Отметим, что если T -

период, то $2T$, $3T$, $4T$, ... - тоже периоды. Из всех периодов будем рассматривать и оценивать наименьший.

Если $T = T_0$ - истинный период (или кратный ему) и погрешности $e(t)$ отсутствуют, то все введенные в предыдущем абзаце функции совпадают между собой и с периодической составляющей: $x(t) = y_1(t) = y_2(t) = y_3(t) = ...$ при всех t из $[0; T]$. При наличии погрешностей полного совпадения не будет. Однако отклонения определяются лишь шумами в различные моменты времени. При этом в качестве оценки периодической составляющей $x(t)$ естественно взять среднее арифметическое $y_{ср}(t)$ функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... (могут быть использованы и другие виды средних величин).

Если же T отличается от истинного периода T_0 (и кратных ему величин), то различия функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ... между собой определяются также и различием значений $x(t)$ в точках, отстоящих друг от друга на интервалы, длина которых кратна T.

В предположении отсутствия погрешностей (т.е. когда $e(t)$ тождественно равно 0) рассмотрим поведение функции $y_{ср}(t)$ на отрезке $[0; T]$ при росте длины интервала наблюдения сигнала A, а потому и при росте числа периодов - целой части числа A/T. Если $T = T_0$ или T кратно T_0, то, как уже сказано, $y_{ср}(t)$ совпадает с периодической составляющей $x(t)$. Если число T/T_0 иррационально, то можно показать, что значения $t + mT \pmod{T_0}$, где m - натуральные числа такие, что $t + mT < A$, асимптотически (при росте A) равномерно заполняют отрезок $[0; T_0]$, а потому при выполнении соответствующих условий регулярности, например, непрерывности периодической составляющей сигнала, функция $y_{ср}(t)$ приближается к константе - среднему значению периодического сигнала $x(t)$, т.е. интегралу от $x(t)$ по отрезку $[0; T_0]$, деленному на T_0. При этом при конечных A функция $y_{ср}(t)$ отлична от константы. Если же число T/T_0 рационально, то наблюдаем промежуточный случай по сравнению с двумя описанными выше, в котором $y_{ср}(t)$, как можно показать, приближается к периодической функции с периодом $T = T_0/n$ при некотором натуральном n. Эта функция получена усреднением n последовательных участков длины T_0/n периодического сигнала $x(t)$. Она не является константой, хотя разброс ее значений меньше, чем для исходного периодического сигнала, поскольку T_0 - минимальная длина периода.

Из сказанного вытекает, что для оценивания T целесообразно ввести два показателя: показатель разброса $F(T; Y) = F(T; y_1(t), y_2(t), y_3(t), ...)$ множества функций $\{y_1(t), y_2(t), y_3(t), ... \}$ на отрезке $[0; T]$ и показатель размаха $G(T; Y) = G(T, y_{ср}(t))$ функции $y_{ср}(t)$ на отрезке $[0; T]$. (Символ Y означает здесь, что показатели разброса и размаха строятся по функции $y(t)$.) При этом показатель разброса нацелен на оценку различий в значениях семейства функций при одном и том же значении аргумента, а показатель размаха - на различие значений одной и той же функции при различных значениях аргумента. Ниже выписан ряд формул для этих показателей в случае непрерывного времени. Для дискретного времени их можно адаптировать двумя способами: либо заменив sup на max, а интеграл на сумму; либо расширив область определения используемых функций на весь отрезок, например, соединив соседние точки отрезками или использовав для заполнения пропусков сплайны более высокого порядка.

В качестве оценки длины периода по фиксированным показателям разброса $F(T; Y)$ и размаха $G(T; Y)$ представляется рациональным использовать то T, при котором отношение $F(T; Y) / G(T; Y)$ впервые (при росте T начиная с 0) достигает минимума. Поскольку показатели разброса $F(T; Y)$ и размаха $G(T; Y)$ могут быть выбраны многими разными способами, можно указанным выше способом построить целое семейство алгоритмов оценивания длины периода, с каждым из которых может быть связано семейство методов оценивания периодической составляющей путем того или иного способа усреднения функций $y_1(t)$, $y_2(t)$, $y_3(t)$, ...

3. Показатели разброса и размаха

Ввести показатели разброса $F(T; Y)=F(T; y_1(t), y_2(t), y_3(t), ...)$ можно разными способами. Пусть $k = [A/T]$. Можно использовать различные функционалы супремумного типа (здесь и далее число слагаемых k не будем указывать в обозначении функционалов). Первым рассмотрим максимальный разброс непосредственно между значениями функций:

$$F_1(T,Y) = \sup\{|\, y_i(t) - y_j(t)\,|, i, j = 1,2,...,k, 0 \le t \le T\}.$$

Второй функционал супремумного типа будет учитывать не произвольные отклонения, а только отклонения от «средней функции», т.е. иметь вид

$$F_2(T,Y) = \sup\{|\, y_i(t) - y_{cp}(t)\,|, i = 1,2,...,k, 0 \le t \le T\}.$$

Третий функционал показывает, какую зону «заметают» значения функций:

$$F_3(T,Y) = \sup\{y_i(t), i = 1,2,...,k, 0 \le t \le T\} - \inf\{y_i(t), i = 1,2,...,k, 0 \le t \le T\}.$$

Для применения функционалов интегрального типа целесообразно сделать замену переменной $q = t/T$ и перейти к функциям $Y_i(q) = y_i(t) = y_i(qT)$, $i = 1, 2, ..., k$, $Y_{cp}(q) = y_{cp}(t) = y_{cp}(qT)$, определенным на отрезке [0; 1]. В качестве показателя разброса представляется полезным рассмотреть то или иное отклонение совокупности функций $Y_i(q)$, $i = 1, 2, ..., k$, друг относительно друга. Можно сказать, что эти функции заполняют некую «трубку», которая тоньше всего при истинном значении периода T, а внутри нее проходит периодическая составляющая $X(q) = x(t) = x(qT)$. Естественно рассмотреть различные функционалы интегрального типа.

Например, можно проинтегрировать максимум модулей попарных разностей:

$$F_4(T,Y) = \int_0^1 \max\{|Y_i(q) - Y_j(q)|, i, j = 1,2,...,k\} dq.$$

Вместо максимума можно проинтегрировать сумму:

$$F_5(T,Y) = \int_0^1 \sum_{i,j=1}^{k} |Y_i(q) - Y_j(q)| dq.$$

Как и для функционалов супремумного типа, естественно рассмотреть показатели разброса относительно «средней функции»:

$$F_6(T,Y) = \int_0^1 \max\{|Y_i(q) - Y_{cp}(q)|, i = 1,2,...,k\} dq,$$

$$F_7(T,Y) = \int_0^1 \sum_{i=1}^{k} |Y_i(q) - Y_{cp}(q)| dq.$$

Следующие четыре функционала, используемые как показатели разброса, аналогичны четырем предыдущим, но включают в себя расчет квадратов некоторых функций:

$$F_8(T,Y) = \int_0^1 [\max\{|Y_i(q) - Y_j(q)|, i, j = 1,2,...,k\}]^2 dq,$$

$$F_9(T,Y) = \int_0^1 \sum_{i,j=1}^{k} \{Y_i(q) - Y_j(q)\}^2 dq,$$

$$F_6(T,Y) = \int_0^1 [\max\{|Y_i(q) - Y_{cp}(q)|, i = 1,2,...,k\}]^2 dq,$$

$$F_7(T,Y) = \int_0^1 \sum_{i=1}^{k} \{Y_i(q) - Y_{cp}(q)\}^2 dq.$$

Список показателей разброса можно существенно расширить. В частности, естественно использовать также расстояния в функциональных пространствах L^p при произвольных $p \geq 1$. А для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних величин.

Показатели размаха также можно ввести самыми различными способами. Например, можно рассмотреть такой показатель:

$$G_1(T,Y) = \sup\{|\, y_{cp}(t) - y_{cp}(s)\,|,\ 0 \le t \le T,\ 0 \le s \le T\} =$$

$$= \sup\{y_{cp}(t),\ 0 \le t \le T\} - \inf\{y_{cp}(t),\ 0 \le t \le T\}.$$

Пусть сделана замена переменной $q = t/T$ и осуществлен переход к функции $Y_{cp}(q) = y_{cp}(t) = y_{cp}(qT)$. Возможными показателями размаха являются:

$$G_2(T,Y) = \int_0^1\int_0^1 |\, Y_{cp}(q) - Y_{cp}(r)\,|\, dqdr,$$

$$G_2(T,Y) = \int_0^1\int_0^1 (Y_{cp}(q) - Y_{cp}(r))^2\, dqdr.$$

Введем среднее значение оценки периодической составляющей:

$$Y_{cp} = \int_0^1 Y_{cp}(q)dq.$$

К естественным показателям размаха относятся, например, такие:

$$G_4(T,Y) = \sup\{|\, Y_{cp}(q) - Y_{cp}\,|,\ 0 \le q \le 1\},$$

$$G_5(T,Y) = \int_0^1 |\, Y_{cp}(q) - Y_{cp}\,|\, dq,$$

$$G_6(T,Y) = \int_0^1 (Y_{cp}(q) - Y_{cp})^2\, dq.$$

Список показателей размаха, как и список показателей разброса, можно значительно расширить. В частности, естественно использовать расстояния в функциональных пространствах L^p при произвольном $p \geq 1$. А для оценивания периодической составляющей применять не только среднее арифметическое, но и другие виды средних - медиану, среднее геометрическое и др. Вопрос о выборе наилучших (в каком-либо смысле) показателей размаха и разброса здесь не обсуждается. Некоторые из причин этого отказа от оптимизации системы показателей рассмотрены ниже.

4. Алгоритмы оценивания

С прикладной точки зрения остается численно минимизировать один или несколько из 66 описанных выше функционалов $F_i(T;\ Y)/G_j(T;\ Y)$, $i = 1, 2, ..., 11$, $j = 1, 2, ..., 6$.

Численная минимизация по одному параметру (возможной длине периода) для современных компьютеров не вызывает проблем, даже если попросту перебирать возможные значения периода, например, с шагом 0,001. По нескольким реальным или смоделированным сигналам можно установить, какой из функционалов позволяет оценить период и периодическую составляющую реально встречающихся сигналов наиболее точно. Возможно и одновременное использование всех или части функционалов, что в соответствии с методологией устойчивости [10 - 13] позволяет установить чувствительность оценок к выбору метода оценивания, найти интервал их разброса. Проведенные в Институте высоких статистических технологий и эконометрики расчеты по реальным и смоделированным данным о временных рядах показали, что описанные выше алгоритмы позволяют оценивать длину периода и восстанавливать периодическую составляющую временного ряда достаточно точно с практической точки зрения.

В литературе по временным рядам проблеме оценивания периода не уделяется необходимого внимания [14 - 29]. Фактически рекомендуют пользоваться либо периодограммой, либо автокорреляционной функцией. С помощью периодограммы (несостоятельной оценки спектральной плотности) можно выделить лишь синусоидальные составляющие, в то время как в кратко рассмотренных выше прикладных задачах периодическая составляющая представляет интерес сама по себе, без разложения на гармоники. Вторая рекомендация более полезна. В качестве оценки периода можно взять наименьшее положительное число, в котором достигается локальный максимум автокорреляционной функции. Эмпирический коэффициент автокорреляции - еще один функционал типа тех, что перечислены выше.

При поверхностном взгляде на проблемы статистического оценивания, как и на иные проблемы прикладной математики, часто возникает желание обсудить «оптимальность» тех или иных процедур. При более глубоком анализе становятся очевидными два обстоятельства. Во-первых, оптимальность имеет быть лишь в рамках той или иной теоретической модели, при отклонениях от которой оптимальность оценки, как правило, пропадает. Например, выборочное среднее арифметическое как оценка математического ожидания случайной величины оптимальна тогда и только тогда, когда распределение результатов наблюдений - гауссово (доказательство этого утверждения приведено в монографии [30]). С другой стороны, для практически любой статистической процедуры можно подобрать свойство оптимальности так, чтобы эта процедура оказалась оптимальной (как подобрать - это уже дело профессионала). Так, например, метод наименьших модулей оптимален, если погрешности имеют распределение Лапласа, а метод наименьших квадратов - когда их распределение гауссово. Поскольку реальные распределения - не Лапласа и не Гаусса, то указанные математические результаты не могут иметь большого практического значения.

Однако представляется полезным получить доказательства состоятельности оценок изучаемых параметров в возможно более широких, например, непараметрических, постановках. Хотя на основе самого факта сходимости нельзя оценить близость оценок к интересующим исследователя параметрам, но получение доказательства состоятельности - первый шаг при изучении скорости сходимости (см. подробнее [31], а также [1, 10, 32]).

5. Состоятельность оценок

Наиболее общий подход к установлению асимптотического поведения решений экстремальных статистических задач развит в статистике нечисловых данных для случая пространств произвольной природы [33, 34]. Согласно этому подходу сначала при фиксированном T доказывается сходимость (по вероятности) при $A \to \infty$ значений функционала (показателя разброса) к некоторой предельной функции, а затем проверяются условия, обеспечивающие сходимость *Argmin* допредельного случайного процесса к *Argmin* этой детерминированной функции.

Свойства алгоритмов приходится изучать в рамках тех или иных вероятностно-статистических моделей. Моделей может быть много. Достаточно вспомнить историю Центральной Предельной Теоремы (ЦПТ) теории вероятностей. Она на протяжении более 200 лет доказывалась во все более и более широких условиях, вплоть до необходимых и достаточных условий Линдеберга - Феллера (после чего начались обобщения на зависимые слагаемые, на суммы случайных элементов гильбертовых пространств и др.). Иногда математические модели далеко выходят за пределы, достаточные для обоснования алгоритмов анализа реальных данных. Так, почти всегда распределения реальных величин дискретны и финитны (сосредоточены на конечном отрезке), а потому, в частности, существуют все моменты. Однако условия финитности и дискретности в вероятностно-статистических моделях часто необоснованно ослабляются. В результате возникают проблемы, не имеющие отношения к реальным данным, например, связанные с измеримостью относительно тех или иных сигма-алгебр. Поэтому ограничимся здесь наиболее простыми моделями из адекватных реальным постановкам. Считаем, что читатель знаком с основными определениями, относящимися к теории случайных процессов.

Теорема 1. Пусть случайный процесс $e(t)$ имеет нулевое математическое ожидание, является стационарным и эргодическим (т.е. выполнена теорема Биркгофа - Хинчина) с непрерывными траекториями. Тогда при фиксированном T и $A \to \infty$ имеем

$$\sup\{| E_{cp}(q)|,\ 0 \le q \le 1\} \to 0$$

(сходимость по вероятности), где $E_{cp}(q) = Y_{cp}(q) - X_{cp}(q)$, т.е. $E_{cp}(q)$ - среднее арифметическое погрешностей $e(qT)$, $e(qT+T)$, $e(qT+2T)$,...

Доказательство теоремы 1 проводится стандартными методами теории стационарных временных рядов (с шагом T) с использованием известного условия достаточно быстрого убывания элементов матрицы Лорана по мере удаления от ее главной диагонали (т.е. условия, необходимого и достаточного для справедливости теоремы Биркгофа - Хинчина). С помощью теоремы 1 можно найти асимптотику введенных выше показателей разброса и размаха.

Теорема 2. В предположениях теоремы 1 при фиксированном T и $A \to \infty$ пронормированные показатели разброса $F_i(T;Y)$ для наблюдаемого сигнала $У$ сближаются по распределению с соответствующими положительными случайными величинами $W_i(T, X, \omega)$, зависящими от T, характеристик случайного процесса $e(t)$ и периодической составляющей X, т.е. существуют числовые последовательности $s_i(k)$ такие, что

$$s_i(k)F_i(T,Y) \Rightarrow W_i(T, X, \omega), \quad i = 1,2,...,11.$$

Доказательство теоремы 2 проводится с помощью достаточно трудоемких (в частности, из-за числа функционалов), но стандартных рассуждений. Они относятся к теории случайных процессов как части теории вероятностей. Эти рассуждения посвящены максимумам (не супремумам, т.к. траектории функции $x(t)$ и случайного процесса $e(t)$ непрерывны) случайных процессов и интегралам от них, с использованием принципа инвариантности (см., например, монографию [35]) и ряда результатов теории стационарных случайных процессов (см., например, монографию [23]). Таким образом, пронормированные функционалы разброса асимптотически не зависят от числа слагаемых - в этом и состоит основной смысл теоремы 2.

Теорема 3. В предположениях теоремы 1 при фиксированном T и $A \to \infty$ показатели размаха для наблюдаемого сигнала $У$ сближаются с соответствующими показателями для периодической составляющей X, т.е.

$$G_j(T, Y) - G_j(T, X) \to 0, \quad j = 1,2,...,6.$$

Для доказательства используются стандартные оценки, основанные на виде конкретных функционалов, задающих показатели размаха. В отличие от теоремы 2 предельные показатели детерминированы.

Аналоги теорем 2 и 3 верны также и при использовании (в качестве показателей разброса и размаха) расстояний в функциональных пространствах L^p при произвольном $p \geq 1$. А для оценивания периодической составляющей - не только среднего арифметического, но и других видов средних - медианы, среднего квадратического, среднего геометрического, обобщенных средних по Колмогорову (см., например, [1, 2, 34]).

Теорема 4. Пусть выполнены условия теоремы 1, периодическая составляющая непрерывна и имеет период T_0. Тогда при фиксированном T и $A \to \infty$ показатели разброса (пронормированные) и размаха стремятся к некоторым детерминированным пределам, зависящим только от T и T_0 , т.е.

$$s_i(k)F_i(T,Y) \to F_i(T,T_0), \quad i = 1,2,...,11,$$

$$G_j(T,Y) \to G_j(T,X), \quad j = 1,2,...,6.$$

(сходимость по вероятности), минимум каждой из функций $F_i(T; T_0)$, i=1, 2, ..., 11, и максимум каждой из функций $G_j(T; T_0)$, j=1, 2, ..., 6, достигается при T = T_0 и при T, кратных T_0, причем у показателей разброса $F_i(T; T_0)$ возможны и иные минимумы, а у показателей размаха $G_j(T; T_0)$ других максимумов нет.

Доказательство вытекает из теорем 2 и 3 и свойств усреднения периодической составляющей при росте длины интервала наблюдения сигнала, описанных в начале статьи. Предельные значения функционала разброса $F_i(T; T_0)$, вообще говоря, показывают разброс случайной погрешности, другими словами, не всегда зависят от периодической составляющей, а потому из-за нормировки на единичный отрезок в ряде случаев оказываются константами. Вместе с тем численные эксперименты показывают, что отмеченная сходимость к пределу является сравнительно медленной. И минимизация непосредственно функционалов разброса (без учета показателей размаха) при конкретной длине сигнала позволяет достаточно точно выделить периодическую составляющую из массива реальных данных. Однако описанные выше теоретические результаты заставили отказаться от первоначальной гипотезы о том, что достаточно использовать только показатели разброса, и привели к необходимости скорректировать алгоритмы, введя деление на показатели размаха.

Теорема 5. В предположениях теоремы 4 оценки, являющиеся первыми локальными минимумами при минимизации по T отношений одного из 11 перечисленных выше показателей разброса к одному из 6 показателей размаха, являются состоятельными оценками истинного периода T_0, а функция $y_{ср}(t)$ является состоятельной оценкой периодической составляющей $x(t)$ на отрезке $[0; T_0]$.

Согласно теоремам 1 - 4 установлена сходимость (по вероятности) значений допредельных функционалов к предельным при каждом конкретном T. Поэтому для доказательства сходимости минимумов допредельных функционалов к минимумам предельных можно воспользоваться общей теорией асимптотического поведения решения экстремальных статистических задач. Условие асимптотической равномерной разбиваемости, введенное и использованное в работе [33], выполнено, как можно показать, в силу непрерывности траекторий случайного процесса (непрерывного сглаживания для временного ряда) и его периодической составляющей. Откуда и вытекает заключение теоремы 5, дающей теоретико-статистическое обоснование использованию системы описанных выше эвристических алгоритмов оценивания длины периода и периодической составляющей. При известной или достаточно точно оцененной длине периода сама периодическая составляющая естественным образом оценивается с помощью усреднения перенесенных к началу координат кусков временного ряда, и в силу теоремы 1 эта оценка является состоятельной. Затем для получения оценки математического ожидания сигнала на всей области его определения указанную оценку можно периодически продолжить.

Замечание. При практическом использовании рассматриваемых алгоритмов целесообразно учитывать дополнительные особенности реальных временных рядов. В частности, обратим внимание на неустойчивость супремумов (в смысле, раскрытом в работах [10 - 13]) по отношению к выбросам (резко выделяющимся наблюдениям) сравнительно с функционалами интегрального типа. Бывают ситуации, когда методики или аппаратура, регистрирующие значения реальных временных рядов, могут допускать сбои в отдельные моменты времени. Например, если происходит валютный кризис типа «черного вторника», когда курс доллара по отношению к рублю, строго говоря, не определен, другими словами, с точки зрения экономических агентов одновременно существует масса сильно отличающихся курсов. Аналогичная ситуация бывает и в целом ряде других случаев. Набор подходящих ассоциаций вызывают решения руководства страны об обмене денежных знаков, особенно с дискриминационными составляющими. Во всех подобных ситуациях временные ряды дают резкие выбросы (всплески), которые затем, как правило, сглаживаются. Поэтому целесообразно в качестве показателей разброса и размаха использовать функционалы интегрального типа. Вопросам оценивания длины периода и периодической составляющей посвящены также статья [36] и разделы в монографиях [1, 2, 32].

Литература

1. Орлов А.И. Прикладная статистика. - М.: Экзамен, 2006. - 672 с.
2. Орлов А.И. Эконометрика. Изд. 3-е, переработанное и дополненное. - М.: Экзамен, 2004. – 576 с.
3. Кондратьев Н., Яковец Ю., Абалкин Л. Большие циклы конъюнктуры и теория предвидения. Избранные труды. - М. : Экономика, 2002. - 768 с.
4. Акаев А.А. Современный финансово-экономический кризис в свете теории инновационно-технологического развития экономики и управления инновационным процессом // Системный мониторинг: глобальное и региональное развитие / Ред. Д.А. Халтурина, А.В. Коротаев. - М.: УРСС, 2009. - С. 141–162.
5. Schumpeter Joseph Alois. Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process. — 1-е изд. — New York - London: McGraw-Hill, 1939. Vol. I. — XVI, 448 p.; Vol. II. — IX, 647 p.
6. Туган-Барановский М.И. Периодические промышленные кризисы. История английских кризисов. Общая теория кризисов. — М.: Наука, 1997. — 573 с.
7. Петров В.М., Мажуль Л.А. Цикличность социокультурной сферы и проблемы среднесрочного прогнозирования ее развития // Математическое и компьютерное моделирование в науках о человеке и обществе. Тезисы докладов Всероссийской конференции. - М.: Госуд. ун-т управления, 1999. - С.63-66.
8. Николаев А.В. Структура исторического цикла // Математическое и компьютерное моделирование в науках о человеке и обществе. Тезисы докладов Всероссийской конференции. - М.: Госуд. ун-т управления, 1999. - С.54-54.

9. Носовский Г.В., Фоменко А.Т. Введение в новую хронологию. (Какой сейчас век?). - М.: КРАФТ+ЛЕАН, 1999. - 488 с.

10. Орлов А.И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.

11. Орлов А.И. Устойчивые математические методы и модели // Заводская лаборатория. Диагностика материалов. 2010. Т.76. №3. С.59-67.

12. Орлов А.И. Устойчивые экономико-математические методы и модели. — Saarbrücken, Lambert Academic Publishing, 2011. — 436 с.

13. Орлов А.И. Новый подход к изучению устойчивости выводов в математических моделях // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 100. С. 146-176.

14. Андерсон Т. Статистический анализ временных рядов. - М.: Мир, 1976. - 757 с.

15. Бендат Дж., Пирсол А. Измерение и анализ случайных процессов. - М.: Мир, 1977. - 464 с.

16. Бриллинджер Д. Временные ряды. Обработка данных и теория. - М.: Мир, 1980. - 536 с.

17. Венсель В.В. Интегральная регрессия и корреляция: статистическое моделирование рядов динамики. - М.: Финансы и статистика, 1983. - 224 с.

18. Гренандер У. Случайные процессы и статистические выводы. -М.: ИЛ, 1961. - 168 с.

19. Журбенко И.Г. Спектральный анализ временных рядов. - М.: Изд-во Моск. ун-та, 1982. - 168 с.

20. Журбенко И.Г. Анализ стационарных и однородных случайных систем. - М.: Изд-во Моск. ун-та, 1987. - 240 с.

21. Кендалл М.Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. - М.: Наука, 1976. - 736 с.

22. Кендэл М. Временные ряды. - М.: Финансы и статистика, 1981. - 200 с.

23. Крамер Г., Лидбеттер М. Стационарные случайные процессы. - М.: Мир, 1969. - 400 с.

24. Ковалева Л.Н. Многофакторное прогнозирование на основе рядов динамики.- М.: Финансы и статистика, 1980. - 104 с.

25. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. - М.: Мир, 1982. - 428 с.

26. Рабинер Р., Гоулд Б. Теория и применение цифровой обработки сигналов. - М.: Мир, 1978. - 848 с.

27. Статистический анализ экономических временных рядов и прогнозирование. (Серия "Ученые записки по статистике", тт.22-23.) - М.: Наука, 1973. - 295 с.

28. Хеннан Э. Многомерные временные ряды. - М.: Мир, 1974. - 576 с.

29. Цветков Э.И. Основы теории статистических измерений. - Л.: Энергоатомиздат, 1986. - 256 с.

30. Каган А.М., Линник Ю.В., Рао С.Р. Характеризационные задачи математической статистики. - М.: Наука, 1972. – 656 с.

31. Орлов А.И. Методы оценки близости допредельных и предельных распределений статистик // Заводская лаборатория. 1998. Т.64. № 5. С.64-67.

32. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч.3. Статистические методы анализа данных. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2012. - 624 с.

33. Орлов А.И. Асимптотика решений экстремальных статистических задач // Анализ нечисловых данных в системных исследованиях. Сб. трудов. Вып.10. - М.: ВНИИСИ, 1982. - С.4-12.

34. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч. 1. Нечисловая статистика. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. — 541 с.

35. Биллингсли П. Сходимость вероятностных мер. - М.: Наука, 1977. – 352 с.

36. Орлов А.И. Метод оценивания длины периода и периодической составляющей сигнала // Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. - Пермь: Изд-во Пермского государственного университета, 1999. - С. 38-49.

References

1. Orlov A.I. Prikladnaja statistika. - M.: Jekzamen, 2006. - 672 s.

2. Orlov A.I. Jekonometrika. Izd. 3-e, pererabotannoe i dopolnennoe. - M.: Jekzamen, 2004. – 576 s.

3. Kondrat'ev N., Jakovec Ju., Abalkin L. Bol'shie cikly kon#junktury i teorija predvidenija. Izbrannye trudy. - M. : Jekonomika, 2002. - 768 s.

4. Akaev A.A. Sovremennyj finansovo-jekonomicheskij krizis v svete teorii innovacionno-tehnologicheskogo razvitija jekonomiki i upravlenija innovacionnym processom // Sistemnyj monitoring: global'noe i regional'noe razvitie / Red. D.A. Halturina, A.V. Korotaev. - M.: URSS, 2009. - S. 141–162.

5. Schumpeter Joseph Alois. Business Cycles: A Theoretical, Historical and Statistical Analysis of the Capitalist Process. — 1-e izd. — New York - London: McGraw-Hill, 1939. Vol. I. — XVI, 448 p.; Vol. II. — IX, 647 p.

6. Tugan-Baranovskij M.I. Periodicheskie promyshlennye krizisy. Istorija anglijskih krizisov. Obshhaja teorija krizisov. — M.: Nauka, 1997. — 573 s.

7. Petrov V.M., Mazhul' L.A. Ciklichnost' sociokul'turnoj sfery i problemy srednesrochnogo prognozirovanija ee razvitija // Matematicheskoe i komp'juternoe modelirovanie v naukah o cheloveke i obshhestve. Tezisy dokladov Vserossijskoj konferencii. - M.: Gosud. un-t upravlenija, 1999. - S.63-66.

8. Nikolaev A.V. Struktura istoricheskogo cikla // Matematicheskoe i komp'juternoe modelirovanie v naukah o cheloveke i obshhestve. Tezisy dokladov Vserossijskoj konferencii. - M.: Gosud. un-t upravlenija, 1999. - S.54-54.

9. Nosovskij G.V., Fomenko A.T. Vvedenie v novuju hronologiju. (Kakoj sejchas vek?). - M.: KRAFT+LEAN, 1999. - 488 s.

10. Orlov A.I. Ustojchivost' v social'no-jekonomicheskih modeljah. — M.: Nauka, 1979. — 296 s.

11. Orlov A.I. Ustojchivye matematicheskie metody i modeli // Zavodskaja laboratorija. Diagnostika materialov. 2010. T.76. №3. S.59-67.

12. Orlov A.I. Ustojchivye jekonomiko-matematicheskie metody i modeli. — Saarbrücken, Lambert Academic Publishing, 2011. — 436 s.

13. Orlov A.I. Novyj podhod k izucheniju ustojchivosti vyvodov v matematicheskih modeljah // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 100. S. 146-176.

14. Anderson T. Statisticheskij analiz vremennyh rjadov. - M.: Mir, 1976. - 757 s.

15. Bendat Dzh., Pirsol A. Izmerenie i analiz sluchajnyh processov. - M.: Mir, 1977. - 464 s.

16. Brillindzher D. Vremennye rjady. Obrabotka dannyh i teorija. - M.: Mir, 1980. - 536 s.
17. Vensel' V.V. Integral'naja regressija i korreljacija: statisticheskoe modelirovanie rjadov dinamiki. - M.: Finansy i statistika, 1983. - 224 s.
18. Grenander U. Sluchajnye processy i statisticheskie vyvody. -M.: IL, 1961. - 168 s.
19. Zhurbenko I.G. Spektral'nyj analiz vremennyh rjadov. - M.: Izd-vo Mosk. un-ta, 1982. - 168 s.
20. Zhurbenko I.G. Analiz stacionarnyh i odnorodnyh sluchajnyh sistem. - M.: Izd-vo Mosk. un-ta, 1987. - 240 s.
21. Kendall M.Dzh., St'juart A. Mnogomernyj statisticheskij analiz i vremennye rjady. - M.: Nauka, 1976. - 736 s.
22. Kendjel M. Vremennye rjady. - M.: Finansy i statistika, 1981. - 200 s.
23. Kramer G., Lidbetter M. Stacionarnye sluchajnye processy. - M.: Mir, 1969. - 400 s.
24. Kovaleva L.N. Mnogofaktornoe prognozirovanie na osnove rjadov dinamiki.- M.: Finansy i statistika, 1980. - 104 s.
25. Otnes R., Jenokson L. Prikladnoj analiz vremennyh rjadov. - M.: Mir, 1982. - 428 s.
26. Rabiner R., Gould B. Teorija i primenenie cifrovoj obrabotki signalov. - M.: Mir, 1978. - 848 s.
27. Statisticheskij analiz jekonomicheskih vremennyh rjadov i prognozirovanie. (Serija "Uchenye zapiski po statistike", tt.22-23.) - M.: Nauka, 1973. - 295 s.
28. Hennan Je. Mnogomernye vremennye rjady. - M.: Mir, 1974. - 576 s.
29. Cvetkov Je.I. Osnovy teorii statisticheskih izmerenij. - L.: Jenergoatomizdat, 1986. - 256 s.
30. Kagan A.M., Linnik Ju.V., Rao S.R. Harakterizacionnye zadachi matematicheskoj statistiki. - M.: Nauka, 1972. – 656 s.
31. Orlov A.I. Metody ocenki blizosti dopredel'nyh i predel'nyh raspredelenij statistik // Zavodskaja laboratorija. 1998. T.64. № 5. S.64-67.
32. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch.3. Statisticheskie metody analiza dannyh. - M.: Izd-vo MGTU im. N.Je. Baumana, 2012. - 624 s.
33. Orlov A.I. Asimptotika reshenij jekstremal'nyh statisticheskih zadach // Analiz nechislovyh dannyh v sistemnyh issledovanijah. Sb. trudov. Vyp.10. - M.: VNIISI, 1982. - S.4-12.
34. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch. 1. Nechislovaja statistika. – M.: Izd-vo MGTU im. N.Je. Baumana, 2009. — 541 s.
35. Billingsli P. Shodimost' verojatnostnyh mer. - M.: Nauka, 1977. – 352 s.
36. Orlov A.I. Metod ocenivanija dliny perioda i periodicheskoj sostavljajushhej signala // Statisticheskie metody ocenivanija i proverki gipotez. Mezhvuzovskij sbornik nauchnyh trudov. - Perm': Izd-vo Permskogo gosudarstvennogo universiteta, 1999. - S. 38-49.