УДК 621.372 + 621.391.26:51916

ОБРАБОТКА СИГНАЛОВ В ВИДЕ СУММЫ ДЕТЕРМИНИРОВАННОЙ ФУНКЦИИ И РЕАЛИЗАЦИИ СТАЦИОНАРНОГО СЛУЧАЙНОГО ПРОЦЕССА

А.Н. Рагозин, В.Ф. Тележкин

В статье проведен анализ обработки сигналов в виде суммы не случайной функции и реализации стационарного случайного процесса с использованием моделей авторегрессии – скользящего среднего.

Ключевые слова: сигнал, временной ряд, цифровой фильтр, прогноз, спектральная плотность.

Временной ряд (сигнал) исследуемого объекта представим в виде:

$$z\,(nT)=f\,(nT)+y\,(nT), \tag{1}$$

где T – период наблюдения (интервал дискретизации непрерывного времени) наблюдаемого сигнала, $n=0,1\dots$ – номер наблюдения, компонента $f(t)$ – неслучайная детерминированная функция времени t, $y(t)$ – стохастическая (случайная) компонента.

Неслучайная функция $f(t)$ в общем случае может иметь многокомпонентную структуру и быть представленной линейными или нелинейными (произведение) отношениями между некоторыми базисными функциями. В качестве базисных функций принято выделять тренд (главная тенденция), регулярные колебания [1]. Также в [1] предложена классификация моделей трендов (главных тенденций), по используемым в них базисным функциям – алгебраические функции, дробно-рациональные функции, квазиполиномы и др. Временной ряд (1) носит нестационарный характер. Отмечается, что многие временные ряды, встречающиеся на практике, обнаруживают нестационарный характер, но при этом их можно отнести к однородным нестационарным рядам, то есть выделяемую в (1) стохастическую компоненту $y(t)$ можно отнести к реализации стационарного случайного процесса [2].

В настоящее время для обработки нестационарных временных рядов широко используется ARIMA (АРПСС) – модель (модель авторегрессии – проинтегрированного скользящего среднего) [2]. Имеется развитый инструментарий для практического использования таких моделей, например [3]. В модели ARIMA, для приведения нестационарного ряда $z\ (nT)$ к стационарному ряду $w\ (nT)=\ \nabla^{d}\ z(nT)$, используется разностный оператор со сдвигом назад ∇ (d – порядок разности), при этом, $\nabla(nT)\ =\ z(nT)\text{-}z(nT\text{-}T)$. Полученный прогнозируемый временной ряд $w(nT)$, являющийся реализацией стационарного случайного процесса, аппроксимируется моделью ARMA (АРСС – авторегрессии – скользящего среднего), при этом: $z\ (nT)=(\nabla^{-1})^{d}\, w(nT)$, ∇^{-1} – оператор, обратный ∇.

Относительно компоненты $f(nT)$, определяющей нестационарность ряда $z(nT)$, сделаны достаточно общие предположения, допускающие сложную динамику поведения $f(nT)$ во времени. Поэтому использование оператора ∇^{d} для устранения нестационарного характера ряда $z(nT)$, при общих предположениях относительно $f(nT)$, может оказаться неэффективным. Необходимо отметить, что на практике ряд $y(nT)$ содержит более быстрые изменения уровня во времени по сравнению с динамикой уровня, определяемой рядом $f(nT)$ и при этом устойчиво сопровождает эту динамику во времени. Модель ARMA (АРСС), аппроксимирующая стационарную стохастическую компоненту $y(nT)$ должна соответствовать устойчивому минимально-фазовому каузальному фильтру [4]. Поэтому, в классическом подходе к обработке временных рядов [2], необходимо предварительное устранение нестационарности $f(nT)$. Важно отметить, что на модель ARMA (АРСС), аппроксимирующую неслучайную (детерминированную) нестационарную компоненту $f(nT)$ не накладываются ограничения на ее устойчивость, так как динамика $f(nT)$ может представлять собой расходящуюся динамику уровня во времени. В случае, если предварительный анализ не позволяет представить наблюдаемый временной ряд $z(nT)$ в виде (1), например, $z(nT)$ представляет собой сложную расходящуюся (сходящуюся или с равномерным распределением) динамику уровней во времени, то для целей дальнейшего анализа используется представление временного ряда $z(nT)$ в виде:

$$z(nT)\approx f(nT),\ \text{(нестационарность)}, \qquad (2)$$

или

$$z(nT) \approx y(nT), \text{ (стационарность).} \tag{3}$$

В случае (2), временной ряд $z(nT)$ не может быть сведен к стационарному ряду, так как исключительно определяется сложной нестационарной динамикой, определяемой рядом $f(nT)$.

Если спектры временных рядов $f(nT)$ и $y(nT)$ ортогональны в области частот, то есть их спектры не перекрываются по частоте или перекрываются незначительно, то их можно разделить, представить отдельными временными рядами $f(nT)$ и $y(nT)$. Можно показать, что оператор *d-й* разности ∇^d (используемый в модели ARIMA (АРПСС), является частным случаем КИХ-фильтра порядка $d+1$ (фильтр с конечной импульсной характеристикой длины $d+1$ отсчетов) [4]. Можно сформировать два КИХ-фильтра Φ_1 и Φ_2 для оптимального разделения временного ряда $z(nT)$ на компоненты:

$$f(nT)= \Phi_1[z(nT)], \tag{4}$$

$$y(nT)= \Phi_2 [z(nT)]. \tag{5}$$

В общем случае для анализа компонент (4) и (5), ряда $z(nT)$ (1), необходимо использовать две разные модели ARMA (АРСС) с различными предъявляемыми к ним требованиями. Если обнаружено, что ряд $z(nT)$ определяется исключительно случайной компонентой (5), то на применяемую модель ARMA (АРСС) накладываются ограничения на устойчивость, если ряд $z(nT)$ определяется детерминированной компонентой (4), то таких требований на применяемую модель ARMA (АРСС) не накладывается.

Модель авторегрессии – скользящего среднего. Модель временного ряда, пригодная для аппроксимации многих встречающихся на практике детерминированных и стохастических рядов данных с дискретным временем $z(nT)$, описывается выходом линейной системы, выражаемым линейным разностным уравнением:

$$z(n)=a_1z(n-1)+a_2z(n-2)+\ldots+a_pz(n-p)-b_1x(n-1)-b_2x(n-2)-\ldots-b_qx(n-q), \tag{6}$$

где n – номер наблюдения (период наблюдения T опущен (nT)); $z(n)$ – наблюдаемый ряд; $x(n)$ – входная (ненаблюдаемая) последовательность, $a_1,\ldots, a_p$; $b_1,\ldots, b_q$ – коэффициенты линейного разностного уравнения (6); p, q – значения, определяющие порядок модели АРСС (p, q). Разностное уравнение (6) связывает входную (ненаблюдаемую) последовательность $x(n)$ с выходной (наблюдаемой) последовательностью $z(n)$ – исследуемым рядом данных.

Модель авторегрессии – скользящего среднего для аппроксимации стохастической компоненты *y(t)*. Компонента $y(t)$ в (1) – реализация стационарного случайного процесса. В этом случае, в модели (6), в качестве входной (ненаблюдаемой) последовательности $x(n)$, используются некоррелированные между собой отсчеты дискретного во времени белого шума. Методика идентификации параметров модели (6) приведена в [2, 3].

Необходимо отметить, что прогноз случайной компоненты $z(t)$ (в данном случае в (6) $z(n)=y(n)$), известной к моменту времени t из наблюдаемого ряда (5), на некоторый будущий момент времени $t+\tau$ $(\tau>0)$, является откликом линейной системы (6), на нулевое входное воздействие $(x(n\text{-}q\text{-}\ell)=0,\ \ell\geq 1,\ \tau = \ell T,\ t+\ \tau =nT+\ell T)$, с заданными начальными условиями модели (6), полученными в результате идентификации по наблюдаемым данным (5). В этом случае, в силу налагаемых ограничений на устойчивость модели (6) прогноз случайной компоненты $z(t+\tau)$ $(\tau>0)$ всегда будет затухающим процессом, стремящимся к нулю $(z(t+\tau)\rightarrow 0)$, при $\tau\rightarrow\infty$ представляет собой одну из затухающих во времени реализаций стационарного случайного процесса $y(t)$. При построении прогноза временного ряда $z(nT)$ (1) важно построение различных прогнозных сценариев, определяемых различными реализациями случайного стационарного процесса $y(t)$. Для получения различных реализаций процесса $y(t)$ (1), (5) идентифицированную модель (6) необходимо рассматривать как формирующий фильтр порядка (p,q), на вход которого подаются некоррелированные отсчеты белого гауссовского шума $x(n)$. При этом на выходе фильтра будет формироваться последовательность $y(n)$ – стационарный во времени случайный процесс. В данном случае входная последовательность $x(n)$ неограниченна во времени, поэтому можно получить различные реализации $y(n)$ произвольной длины.

Задачу идентификации модели (6) по наблюдаемой реализации случайной компоненты $y(t)$ (1). (5) можно также рассматривать как задачу аппроксимации спектральной плотности мощности (СПМ) по частоте процесса $y(n)$ СПМ модели (6). В данном случае, генерируемый моделью (6) случайный процесс $y(t)$ будет иметь СПМ модели (6), которая может быть рассчитана по коэффициентам модели [4]. СПМ является важной характеристикой, так как определяет распределение по частоте дисперсии стационарного случайного процесса $y(t)$.

Модель авторегрессии – скользящего среднего для аппроксимации детерминированной компоненты *f(t)*. В результате анализа результатов работ [1, 5–8], можно прийти к выводу, что задача идентификации параметров модели (6), когда $z(n)=f(n)$ – детерминированная (неслучайная) функция, сводится к задаче определения параметров модели (6) по наблюдаемому ряду (1), (4), для случая, когда входной (ненаблюдаемый) ряд $x(n)$ является единичным воздействием, при нулевых начальных условиях разностного уравнения (6). При этом наблюдаемый временной ряд $z(n)=f(n)$ является реакцией идентифицированной линейной системы (6) на единичное воздействие, т.е. импульсной характеристикой (откликом) этой системы. Расчет, в соответствии с (6), импульсного отклика модели (6) ARMA

(АРСС (p,q)), после ее идентификации, за пределы наблюдаемого ряда (1), (4), будет являться прогнозом ряда $z(n)=f(n)$ за пределы его наблюдения.

Как показано в [6, 7, 8], рассчитанные параметры модели (6) дают возможность распознавания конкретного вида математической функции, описывающей зависимость $f(t)$ по наблюдаемому временному ряду $f(nT)$ (1), (4). Более детально этот вопрос рассмотрен, например, в [8].

В заключение необходимо отметить, что обработка сигнала с целью прогноза временного ряда $z(nT)$ (1) складывается из прогнозов рядов $f(nT)$ (2) и $y(nT)$ (3). Прогноз $y(nT)$ в классическом случае [2] задается затухающей во времени реализацией случайного процесса и стремится с течением времени к нулю. С использованием (6) возможно генерирование незатухающего с течением времени стационарного случайного процесса $y(nT)$, что позволит получить множество реализаций $y(nT)$ для построения прогнозных сценариев исследуемого ряда $z(nT)$. Прогноз детерминированной составляющей $f(nT)$ является откликом системы (6) на единичное воздействие. По параметрам модели (6) возможно восстановление вида математической функции, определяющей отклик $f(nT)$. Представление наблюдаемого временного ряда (1) в виде суммы детерминированной (4) и стохастической (5) компонент и параметризация их различными ARMA (АРСС) моделями расширяет возможности обработки с целью прогнозирования и позволяет строить прогнозы более качественно, чем в классическом случае использования одной ARIMA (АРПСС) модели. Если наблюдаемый ряд невозможно разделить на компоненты (4), (5), то он относится или к детерминированному нестационарному процессу (2) или к реализации стационарного случайного процесса (3). В этом случае для целей прогнозирования используется одна из рассмотренных в работе моделей ARMA (АРСС).

Библиографический список

1. Семёнычев, В.К. Идентификация экономической динамики на основе моделей авторегрессии / В.К. Семёнычев. – Самара: АНО «Изд-во СНЦ РАН». – 2004. – 243 с.
2. Бокс, Дж. Анализ временных рядов. Прогноз и управление / Дж. Бокс, Г.М. Дженкинс. – М.: Мир, 1974. – Вып. 1, 2.
3. Боровиков, В.П. Прогнозирование в системе STATISTICA в среде Windows. Основы теории и интенсивная практика на компьютере: учебное пособие / В.П. Боровиков, Г.И. Ивченко. – М.: Финансы и статистика, 1999. – 384 с.
4. Рабинер, Л. Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978. – 848 с.
5. Лернер, З.Д. Алгоритм идентификации линейных стационарных систем и распознавания сигналов / З.Д. Лернер // Изв. вузов. Радиоэлектроника. – 1994. – № 4. – С. 74–76.

6. Никитин, Д.А. Синтез рекурсивных цифровых фильтров по импульсной характеристике, определяемой элементарной математической функцией / Д.А. Никитин, В.Х. Ханов // Цифровая обработка сигналов. – 2008. – № 3. – С. 10–14.

7. Кохрейн, Дж. Прогнозирование и импульсные отклики в линейных системах / Дж. Кохрейн // Квантиль. – 2006. – № 1. – С. 21–26.

8. Ханов, В.Х. Алгоритм анализа числовых последовательностей / В.Х. Ханов, Д.А. Никитин // Вестник СибГАУ. – 2006. – № 6 (13). – С. 11–15.

К содержанию