

МЕТОД УЗКОПОЛОСНОГО СПЕКТРАЛЬНОГО АНАЛИЗА В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Филимонов А.Б., Рашиди А.Х.

Московский авиационный институт (НИУ), г. Москва

Ключевые слова: временные ряды, колебания, прогнозирование, узкополосный спектральный анализ.

Аннотация. Теоретический и практический интерес представляет разработка непараметрических методов анализа случайных колебаний временных рядов. В докладе обсуждается концепция прогнозирования временных рядов на основе методологии узкополосного спектрального анализа нерегулярных колебаний уровней ряда.

Во многих задачах прикладной статистики объектом исследований являются задачи прогнозирования временных рядов с явно выраженной колебательной структурой. Работоспособность и условия применимости методов и алгоритмов прогнозирования определяются рядом факторов - имеющимися априорными данными, заданным горизонтом прогнозирования, требованиями к точности и робастности формируемых прогностических оценок.

Если временные ряды имеют периодическую и квазипериодическую составляющие, то для их изучения может применяться спектральный подход - методы гармонического и периодограммного анализа [1, 2]. Так, посредством Фурье-анализа выделяются сезонная и циклическая компоненты экономического временного ряда [3]. Вместе с тем, динамическая структура временного ряда может включать непериодические и случайные колебания с различными временными масштабами. В этом случае необходимы дополнительные теоретические инструменты анализа временных рядов, в число которых целесообразно включать концепцию узкополосных случайных процессов [4].

Рассмотрим временной ряд $y(t)$, где $t \in \mathbf{Z}_+$ - дискретное время (порядковый номер наблюдения), $y \in \mathbf{R}$ - уровень ряда (результат наблюдения); $\mathbf{Z}_+$ - множество неотрицательных целых чисел, $\mathbf{R}$ - поле вещественных чисел.

Положим, прогноз уровней конкретного временного ряда строится по данным его наблюдения в течение отрезка времени $[0, T] \subset \mathbf{Z}_+$ - *базы прогноза* [5]. Соответствующую статистическую оценку среднего уровня ряда в некоторый момент времени $\theta > T$ будем обозначать через $\hat{y}(\theta|T)$. Для h -шагового прогноза имеем: $\theta = T + h$.

В общем случае временной ряд является нестационарным и его можно разложить на четыре составляющие:

$$y(t) = y_d(t) + y_p(t) + z(t) + \varepsilon(t), \quad (1)$$

где $y_d(t)$ - тренд, $y_p(t)$ - периодическая составляющая, $z(t)$ - стационарные случайные колебания, $\varepsilon(t)$ - шум, который считаем быстро осциллирующим случайным процессом с нулевым средним.

Формирование прогноза основано на последовательном выделении и экстраполяции всех трех компонент ряда: $y_d(t)$, $y_p(t)$ и $z(t)$.

Тренд является детерминированной непериодической составляющей процесса $y(t)$, описывающей долгосрочную тенденцию его изменения.

Для построения и многошаговой экстраполяции тренда $y_d(t)$ могут быть использованы методы аналитического выравнивания, основанные на постулировании динамической структуры тренда и выделении трендовой составляющей ряда методом наименьших квадратов (МНК). Их слабым местом является субъективность выбора типа трендовых моделей. Здесь предварительный выбор подходящего типа тренда можно основывать на данных механического сглаживания ряда.

Положим, $\hat{y}_d(t)$ - результат вычисления тренда. Удаляя его из исходного ряда (1), получаем остаточный ряд $y'(t)$:

$$y'(t) = y(t) - \hat{y}_d(t). \quad (2)$$

Оценку $\hat{y}_p(t)$ периодической компоненты ряда $y_p(t)$ можно основывать на применении спектральных методов анализа к остаточному ряду (2).

Исключение из исходного ряда $y(t)$ полученных оценок трендовой и периодической компонент дает остаточный ряд

$$y''(t) = z(t) + \eta(t).$$

Здесь помеха $\eta(t)$ обусловлена шумом $\varepsilon(t)$, а также ошибками в расчете тренда и периодической компоненты исходного ряда.

Дальнейшие действия направлены на построении модели случайного процесса $z(t)$. Для этой цели применяют параметрические и непараметрические подходы [6].

В случае параметрического подхода задается определенная функциональная или динамическая структура временного ряда с неизвестными параметрами, для оценки которых обычно применяется МНК. Класс весьма популярных параметрических моделей ARMA и ARIMA базируется на аппарате стохастических разностных уравнений и постулировании вероятностных свойств наблюдаемых процессов. Задачи параметрической идентификации таких моделей предполагают статистическую обработку выборок большого объема. Основным недостатком параметрического подхода - неробастность получаемых решений.

Теоретический и практический интерес представляет разработка непараметрических методов анализа случайных колебаний временных рядов, способных составить хорошую альтернативу параметрическим методам. Здесь весьма полезной может оказаться методология узкополосного спектрального анализа.

При исследовании процесса $z(t)$ будем представлять его в виде аддитивной смеси узкополосных случайных сигналов.

Стационарный в широком смысле случайный процесс называется *узкополосным*, если его спектральная плотность мощности сосредоточена в относительно узкой полосе частот около некоторой фиксированной частоты ω_0 . Условие узкополосности выражается неравенством:

$$\Delta\omega/\omega_0 \ll 1,$$

где $\Delta\omega$ - ширина полосы спектра.

Основопологающая идея исследования узкополосных процессов (УП) заключается в том, чтобы рассматривать их как колебания с амплитудной и частотной модуляцией:

$$x(t) = A(t) \cos(\omega_0 t + \varphi(t)), \quad (3)$$

где $A(t)$ - амплитуда (амплитудная огибающая), $\varphi(t)$ - фаза (фазовая функция), которые медленно меняются по отношению к гармоническому колебанию с частотой ω_0 .

Для определения статистических характеристик огибающей и фазы УП удобно представлять его в комплексной форме - в виде так называемого *аналитического* сигнала [5]:

$$x_a(t) = x(t) + j x_{\perp}(t).$$

Его вещественная часть совпадает с исходным сигналом (3):

$$\operatorname{Re}[x_a(t)] = x(t) = A(t) \cos(\omega_0 t + \varphi(t)),$$

а его мнимая часть

$$\operatorname{Im}[x_a(t)] = x_{\perp}(t) = A(t) \sin(\omega_0 t + \varphi(t))$$

является сопряженным к исходному сигналу, т.е. определяется преобразованием Гильберта

$$\hat{x}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau) d\tau}{t - \tau}. \quad (4)$$

Из выражения (4) следует, что преобразование Гильберта есть результат свертки функции $x(t)$ с функцией

$$h(t) = \frac{1}{\pi t},$$

называемой ядром преобразования Гильберта. Последнее можно интерпретировать как импульсную характеристику линейного фильтра, на выходе которого формируется ортогональное дополнение входного сигнала.

Комплексная частотная характеристики такого фильтра равна

$$K(j\omega) = -j \operatorname{sign}(\omega),$$

где sign - функция знака:

$$\operatorname{sign}(\omega) = \begin{cases} 1, & \omega > 0; \\ 0, & \omega = 0; \\ -1, & \omega < 0. \end{cases}$$

Таким образом, прямое преобразование Гильберта можно рассматривать как результат прохождения сигнала $x(t)$ через линейный четырехполюсник, сдвигающий фазу всех гармонических составляющих спектра на угол $-\pi/2$.

Процедура определения амплитудной огибающей и фазовой функции состоит из следующих шагов. Сначала получаем гильберт-образ (4) исходного сигнала $x(t)$. Далее определяется огибающая $A(t)$ как модуль аналитического сигнала:

$$A(t) = \sqrt{x^2(t) + x_{\perp}^2(t)}.$$

Полная фаза определяется как аргумент аналитического сигнала:

$$\psi(t) = \arg(x_a(t)) = \arccos(x(t)/A(t)).$$

Введем обозначение: если F - частотный диапазон сигнала $z(t)$, то для любого его подмножества $\tilde{F} \subset F$ через $z(\tilde{F}, t)$ будем обозначать сигнал, образованный суммированием всех гармоник сигнала $z(t)$ с частотами $f \in \tilde{F}$.

Пусть частотный диапазон сигнала $z(t)$ разбит на M полос:

$$F = \bigcup_{i=1}^M F_i. \quad (5)$$

Выполним спектральное разложение

$$z(t) = \sum_{i=1}^M z(F_i, t). \quad (6)$$

Спектр слагаемого $z(F_i, t)$ сосредоточен в частотном поддиапазоне F_i . Выделение указанных спектральных компонент исходного сигнала возможно осуществлять посредством соответствующего набора полосовых фильтров.

Введем обозначения: f_i - центральная точка интервала F_i и

$$\omega_i = 2\pi f_i \quad (i = \overline{1, M}).$$

Положим, все частотные диапазоны F_i в (5) удовлетворяют условию узкополосности. Тогда сигнал $z(F_i, t)$ является узкополосным и его можно представить в виде квазигармонического колебания

$$z(F_i, t) = A_i(t) \cos(\omega_i t + \phi_i(t)) \quad (7)$$

со слабой нестационарностью амплитуды и начальной фазы.

Применение узкополосного спектрального анализа (на отрезке времени $[0, T]$) к каждой из спектральных компонент в разложении (6) позволяет получить оценки $\hat{A}_i(t)$ и $\hat{\phi}_i(t)$ параметров колебаний (7). В тоге будем располагать моделями этих колебаний:

$$\hat{z}(F_i, t) = \hat{A}_i(t) \cos(\omega_i t + \hat{\phi}_i(t)). \quad (8)$$

Согласно (9) прогнозирование процесса $z(t)$ сводится к экстраполяции квазигармонических колебаний (8) на будущие моменты времени $\theta > T$:

$$z(\theta | T) = \sum_{i=1}^M \hat{z}(F_i, \theta). \quad (9)$$

Отметим одну важную особенность динамической структуры прогноза, формируемого согласно описанной схеме узкополосного спектрального анализа колебаний уровней временного ряда $z(t)$. Длиннопериодические волны в его спектральном составе (6) в меньшей степени подвержены неконтролируемым изменениям в течение интервала прогнозирования по сравнению с короткопериодическими колебаниями, и это обстоятельство следует учитывать при обосновании задания горизонта прогноза и оценке его динамической точности.

Список литературы

1. Pollock D.S.G. A Handbook of Time Series Analysis, Signal Processing, and Dynamics. - San Diego, Calif.; London: Academic Press, 1999. - 782 p.
2. Бриллинджер Д. Временные ряды. Обработка данных и теория. - М.: Мир, 1975. - 536 с.
3. Кендэл М. Временные ряды. - М.: Финансы и статистика, 1981. - 199 с.
4. Сергиенко А.Б. Цифровая обработка сигналов. - СПб.: Питер, 2002. - 608 с.
5. Льюис К.Д. Методы прогнозирования экономических показателей. - М.: Финансы и статистика, 1986. - 133 с.
6. Большаков А.А., Каримов Р.Н. Методы обработки многомерных данных и временных рядов. - М.: Горячая линия–Телеком, 2007. - 522 с.

Сведения об авторах:

Филимонов Александр Борисович – д.т.н., с.н.с., профессор, МАИ, г.Москва;

Рашиди Алашти Ходжатоллах – аспирант, МАИ, г.Москва.

THE METHOD OF NARROW-BAND SPECTRAL ANALYSIS IN TASKS OF TIME SERIES FORECASTING

Filimonov A.B., Rashidi A.H.

Keywords: time series, oscillations, forecasting, narrow-band spectral analysis.

Abstract. The development of nonparametric methods of the analysis of random oscillations of time series represents the theoretical and practical interest. The conception of time series forecasting based on the methodology of narrow-band spectra analysis of irregular fluctuations of series levels is discussed in the report.

References

1. Pollock D.S.G. A Handbook of Time Series Analysis, Signal Processing, and Dynamics. - San Diego, Calif.; London: Academic Press, 1999. - 782 p.
2. Brillinger, D. Time Series. Data Processing and Theory. - M.: Mir, 1975. - 536 p.
3. Kendell M. Time Series. - Moscow: Finance and statistics, 1981. - 199 p.
4. Sergienko A. Digital Signal Processing. - SPb.: Peter, 2002. - 608 p.
5. Lewis, K.D. Methods of Forecasting Economic Indicators. - Moscow: Finance and statistics, 1986. - 133 p.
6. Bolshakov A.A., Karimov R.N. Methods of Processing of Multidimensional Data and Time Series. - M.: Hot line–Telecom, 2007. - 522 p.