ПРИМЕНИЕ ВЕЙВЛЕТ-АНАЛИЗА ДЛЯ ПРОГНОЗИРОВАНИЯ СТАЦИОНАРНЫХ ВРЕМЕННЫХ РЯДОВ

Лобач С. В.
Кафедра математического моделирования и анализа данных,
Факультет прикладной математики и информатики,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: lobachЫ@bsu.by

Рассматривается задача непараметрического прогнозирования временных рядов на основе дискретного вейвлет-преобразования. Построены прогнозирующие статистики в виде разложения по вейвлет-базису.

Введение

Важной проблемой в теории временных рядов является задача прогнозирования будущих значений временного ряда по прошлым данным [1]. Пусть $X_T=\{x(t)\}$, t=$0,1,\ldots,T-1$, – наблюдаемый временной ряд, математическая модель которого соответствует некоторой заданной гипотетической модели. Требуется построить прогнозирующую статистику $g(X_T)$ для оценивания значения временного ряда в момент времени $t = T-1+\tau$, τ – горизонт прогнозирования. В зависимости от типа прогноза рассматривают точечный и интервальный прогноз. В случае точечного прогноза необходимо построить прогнозирующую статистику $g(X_T)$, т.е. $\hat{x}(T-1+\tau) = g(X_T)$, а в случае интервального прогноза – две прогнозирующих статистики $g_1(X_T)$ и $g_1(X) < g_2(X)$, таких, что с вероятностью $\gamma \in (0,1)$ значение $x_{T-1+\tau}$ попадает в интервал $(g_1(X_T), g_2(X_T))$. В зависимости от сделанных предположений относительно гипотетической модели различают праметрическое и непараметрическое прогнозирование. В случае параметрического моделирования постулируется некоторая гипотетическая модель временного ряда (например, регрессионная модель, модель авторегрессии, ARMA-модель, ARIMA-модель, ARCH-модель и т.д.), которая определена с точностью до параметров. По наблюдениям X_T строят оценки параметров модели. Построенная оценка подставляется в уравнения, задающие гипотетическую модель временного ряда. В случае непараметрического прогнозирования предполагается, что модель временного ряда задается в виде $x(t) = f(t)+\varepsilon_t$, где $\{\varepsilon_t\}$ – последовательность случайных величин, $f(x)$ – неизвестная функция. Проблема заключается в построении оценки этой функции. Обычно для построения оценки функции $f(t)$ используются ядерные оценки.

В данной работе рассматривается подход, основанный на использовании вейвлетов для оценки функции $f(t)$.

I. Вейвлеты для временных рядов

Пусть $f(t)$, $t \in [0,T]$, – некоторая функция, причем ее область определения может быть и конечным множеством $\{0,1,\ldots,T-1\}$. Под вейвлет-анализом понимается разложение функции в следующей форме [2].

$$f(t) = \sum_j \sum_k c_{jk}\psi_{jk}(t), \qquad (1)$$

где

$$\psi_{jk}(t) = 2^{j/2}\psi\left(2^j t - k\right). \qquad (2)$$

Функция $\psi(t)$ в (2) называется материнским вейвлетом. Часто в приложениях в качестве $\psi(t)$ берут функцию Хаара:

$$\psi(t) = \begin{cases} 1, & 0 \le t < \dfrac{T}{2}, \\ -1, & \dfrac{T}{2} \le t < T. \end{cases}$$

В случае непрерывного времени и ортогональности свойства функций $\psi_{jk}(t)$ коэффициенты разложения (1) вычисляются по формулам

$$c_{jk} = \frac{\int\limits_0^T \psi_{jk}(t) f(t)dt}{\int\limits_0^T \psi_{jk}(t)^2 dt}. \qquad (3)$$

Формула (3) отражает процедуру анализа функции, а формула (1) – процедуру синтеза функции.

Для непараметрической модели временного ряда

$$x(t) = f(t) + \varepsilon(t)$$

вейвлет-анализ заключается в построении оценок коэффициентов $\hat{c}_{jk}$ и в ограничении количества членов в разложении (1).

II. Статистические свойства конечного вейвлет-преобразования временного ряда

Рассмотрим последовательность $x(t)$, $t = 0,1,\ldots,T-1$, где $\{x(t), t \in Z\}$ является стационарным случайным процессом с нулевым математическим ожиданием ковариационной функции $\gamma(t)$, и $T = 2^M$. Для $j = 1,2,\ldots,M$

и $k = 0, 1, \ldots, \left(2^{(M-J)} - 1\right)$, определим конечное вейвлет-преобразование заданной последовательности наблюдений

$$d_{jk} = \sum_{t=0}^{T-1} x(t)\psi_{jk}(t).$$

Непосредственными вычислениями можно убедиться в том, что

$$E\{d_{jk}\} = 0,$$

$$\mathrm{Var}\{c_{jk}\} = \sum_{l=-(T-1)}^{T-1} \gamma(u) \sum_{t=0}^{T-1-|l|} \psi_{jk}\psi_{jk}(t+|l|).$$

Вейвлеты Хаара являются наиболее широко используемыми вейвлетами из-за простоты их компьютерной реализации. По сути применяются две операции: взятие среднегоарифметического значения и взятие полуразности между двумясоседними значениями функции. Проиллюстрируем это на примере. Пусть задана дискретная функция $f(x) = \{7\ 5\ 1\ 9\}$.

На уровне разрешения $j = 4$ имеем все значения дискретной функции. На уровне $j = 2$ значения 6 и 5 получаются как средние значений 7 и 5, 1 и 9; детализирующие коэффициенты -1 4 получаются следующим образом: $(5-7)/2$, $(9-1)/2$. Этот процесс продолжается до тех пор, пока не будет достигнут уровень разрешения $j = 1$. Вейвлет-проеобразование Хаара функции имеет вид $H\{f(t)\} = \{5.5\ -0.5\ -1\ 4\}$, т.е.получается комбинациейобщего среднегозначения и детализирующих коэффициентов.

В случае вейвлета Хаара коэффициенты d_{jk} принимают вид

$$d_{jk} = 2^{-j/2}\left(\sum_{t=k2^j}^{2^j(k+0.5)} x(t) - \sum_{t=2^j(k+0.5)}^{2^j(k+1)} x(t)\right).$$

В качестве порогового значения будем использовать универсальное пороговое значение для всех уровней разложения j, предложенное в работе [3],

$$\lambda = \sigma \cdot \sqrt{2\log_2 T},$$

где σ – среднеквадратическое отклонение последовательности $x(t)$. Таким образом, коэффициенты разложения (1) будут определяться по формулам

$$d^*_{jk} = \begin{cases} d_{jk}, & \text{если} \quad |d_{jk}| > \lambda, \\ 0, & \text{иначе.} \end{cases}$$

III. Компьютерные эксперименты

В качестве временного ряда использовались данные статистики торгов иностранными валютами за $T = 2^7 = 128$ торговых дней в период с 01.09.26 по 01.03.17. Получились следующие результаты: 1) доллар США – реальное значение 1.9002 на 01.03.17; спрогнозированное 1.9007; 2) евро – реальное значение 2.0003; спрогнозированное 1.9995. Результаты компьютерного моделирования показывают достаточно высокую эффективность использования вейвлетов в статистическом прогнозировании.

1. Харин, Ю. С. Оптимальность и робастность в статистическом прогнозировании / Ю. С. Харин. – Минск: БГУ, 2008. – 263 с.
2. Смоленцев, Н. К. Введение в теорию вейвлетов / Н. К. Смоленцев. – М.-Ижевск: НИЦ “Регулярная и хаотическая динамика”, 2010. – 292 с.
3. Donoho, D. L. Adapting to unknown smoothness via wavelet shrinkage / D. L. Donoho, I. M. Johnstone // Journal of the American Statistical Association. – 1995. – Vol. 90. – P. 1200–1221.