

c, b $/$ d

{

—

— /

—

,

 $a \quad d$

,

,

,

УДК 519.22

Е.А. МИРОТИН

ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ С ИСПОЛЬЗОВАНИЕМ НЕПРОРЕЖЕННОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И ОДНОСТОРОННИХ ФИЛЬТРОВ

A family of algorithms using non-decimated wavelet transform for time series forecasting is covered. The use of one-sided (backward) filters is grounded. Filter choosing criterion is put, and an example of a filter sufficing the above criterion in case of polynomial trend is constructed.

В последнее время на практике все более широко используются алгоритмы прогнозирования временных рядов, основанные на вейвлет-преобразовании, в частности на так называемом непрореженном вейвлет-преобразовании (НПВП) [1–5]. Подобные алгоритмы, как правило, состоят из разложе-

ния исходного временного ряда в совокупность последовательностей вейвлет-коэффициентов, построения прогнозов для каждой такой последовательности и дальнейшего восстановления значений исходного временного ряда и спрогнозированных значений по имеющимся и спрогнозированным значениям коэффициентов.

Общим свойством таких алгоритмов является то, что к детализирующим вейвлет-коэффициентам применяются методы прогнозирования, основанные на известной модели (например, авторегрессии (АР)). При этом авторы [1, 2, 4], как правило, ограничиваются эмпирическими соображениями относительно того, что полученные в результате разложения детализирующие коэффициенты являются, по их выражению, «более стационарными», чем исходный временной ряд. В [6] проведено теоретическое обоснование выбора фильтра и высказано предположение о соответствии детализирующих вейвлет-коэффициентов стационарной модели. Однако на практике оказывается, что предложенный в [6] подход к выбору фильтра имеет значительный недостаток – краевой эффект в конце интервала наблюдений, вносящий существенный тренд в детализирующие коэффициенты. В настоящей работе приводится модификация алгоритма из [6], проведенная за счет нового подхода к выбору фильтра, и рассматривается критерий выбора фильтра.

Математическая модель. Пусть имеются наблюдения временного ряда $Y(t)$ в моменты времени $t = 1, \dots, T$. Для заданного $\tau \in \mathbb{N}$ требуется построить прогноз $\hat{Y}(t)$ значений $Y(t)$ в моменты времени $t = T + 1, \dots, T + \tau$. Непрореженные вейвлет-коэффициенты определяются следующим образом [1]:

1. Задается число уровней разрешения M , на которые производится вейвлет-декомпозиция временного ряда.

2. Строится семейство последовательно «сглаживаемых» временных рядов:

$$c_0(t) = Y(t), t \in \mathbb{Z},$$

$$c_i(t) = \sum_{k=-n}^n \phi(k) c_{i-1}(t + 2^{i-1} k), t \in \mathbb{N}, i = 1, \dots, M, \quad (1)$$

где $\phi(\cdot)$ – фильтр длины $2n + 1$, который, как правило, удовлетворяет условию нормировки $\sum_{k=-n}^n \phi(k) = 1$.

3. Вычисляются разности $w_i(t) = c_{i-1}(t) - c_i(t), t \in \mathbb{Z}$.

Тогда исходный временной ряд можно представить в виде

$$Y(t) = c_M(t) + \sum_{j=1}^M w_j(t).$$

Последнее равенство является простым аддитивным алгоритмом восстановления исходного временного ряда. При этом компонента $c_M(t)$ имеет смысл сглаженного временного ряда и может рассматриваться как оценка тренда, а компоненты $w_j(t), j = 1, \dots, M$, называемые детализирующими, имеют смысл уточняющих составляющих. Чем больше значение M , тем более гладким получается остаток, и чем больше j , тем более крупные детали характеризуются слагаемым w_j .

Очевидно, что на практике t , как правило, изменяется в конечных пределах (например, от 1 до некоторого $T > 1$). Тогда формула (1) некорректна при t , близком к 1 или T . Эта некорректность, как правило, устраняется путем изменения формулы (применяется так называемое уменьшение базы), однако в изученных случаях оказывается, что, например, при быстро растущем тренде (как в случае с полиномиальным [6]) это изменение привносит существенные искажения в модель остатков. Иными словами, хотя в теории в [6] показан метод построения фильтра, обеспечивающий остатки с нулевым математическим ожиданием, на практике за счет краевых эффектов остатки имеют сильно растущий к концу интервала тренд.

Для исследования свойств стационарности в [6] предложен аналог дискретного НПВП для непрерывного случая. Рассмотрим случайный процесс $c(t) = c(t, \omega), t \in \mathbb{R}, \omega \in \Omega$. Пусть имеются число $p > 0$ и функция φ с компактным носителем, такая, что

$$\int_{-\infty}^{\infty} \varphi(t) dt = 1. \quad (2)$$

Определим зависящий от этих параметров оператор сглаживания (фильтрации):

$$W_A c(t) = \int_R c(t + pu) \varphi(u) du. \quad (3)$$

Результатом применения данного оператора является случайный процесс, который будем называть аппроксимирующей составляющей исходного процесса и в дальнейшем обозначать $d(t)$. Определим детализирующий оператор следующим образом:

$$W_D c(t) = c(t) - W_A c(t). \quad (4)$$

Обозначим $w(t) = W_D c(t)$ и определим этот процесс как детализирующую составляющую исходного процесса. Таким образом, непореженное непрерывное вейвлет-преобразование по исходному процессу $c(t)$ строит его аппроксимирующую составляющую $W_A c(t)$ и детализирующую $W_D c(t)$. Обратное преобразование имеет вид

$$c(t) = d(t) + w(t) = W_A c(t) + W_D c(t).$$

Обычное НПВП можно считать дискретизацией данного непрерывного преобразования, при этом процедура производится в M шагов, а в качестве p выбираются последовательные степени 2. На практике аппроксимирующая составляющая считается оценкой тренда процесса, а детализирующая – стохастической компонентой [1, 2].

Стационарность детализирующей составляющей. Рассмотрим математическое ожидание $m(t)$ случайного процесса $c(t)$. Будем полагать, что функция $m(t)$ измерима и локально ограничена на R . В [6] вводится понятие вариативности математического ожидания.

Определение 1. Пусть K – некоторое положительное число. Рассмотрим подпространство констант $C = \{g(t) : g(t) = \text{const}, t \in [-K, K]\}$ в весовом гильбертовом пространстве $L_2([-K, K], \frac{\text{mes}}{2K})$, где mes – мера Лебега на прямой. Вариативностью математического ожидания процесса $c(t)$ размера K назовем величину $N_K^{\text{mean}} \langle c(t) \rangle = \|m(t) - m_C\|_{L_2(R, \frac{1}{2K} I_{[-K, K]})}$ – расстояние от функции математического ожидания этого процесса до подпространства C (здесь $m_C = \frac{1}{2K} \int_{-K}^K m(t) dt$ – проекция $m(t)$ на C). Данная величина используется как числовой показатель нестационарности случайного процесса по математическому ожиданию.

Показано [6], что математическое ожидание детализирующего процесса $w(t)$ имеет вид $n(t) = Bm(t) = m(t) - \int_R m(t+u)h(u)du$, где

$$h(t) = \frac{1}{p} \varphi\left(\frac{t}{p}\right) \quad (5)$$

– нормированный фильтр. В работе [6] вводится также следующее

Определение 2. Будем говорить, что фильтр h выделяет из случайного процесса тренд (на классе математических ожиданий M), если для всякого $m \in M$ для достаточно больших K выполняется $N_K^{\text{mean}} \langle Bm(t) \rangle = 0$.

Для того чтобы к прогнозированию случайного процесса можно было применить описанный подход, выделяемая детализирующая составляющая должна быть стационарной. Одно из необходимых условий – это то, что ее математическое ожидание – константа, т. е. что фильтр h выделял из случайного процесса тренд.

Краевой эффект и применение односторонних (неупреждающих) фильтров. Как уже отмечалось, формула (1) не применима на практике при t , близком к 1 или T . В случае быстро растущего тренда (например, тренд является кубическим полиномом, $T = 100$) даже применение уменьшения базы приводит к искажениям, которые мало заметны в сглаженном сигнале, но хорошо заметны в разности между исходным сигналом и отфильтрованным. Таким образом, детализирующая компонента приобретает сильно растущий к концу тренд, делающий применение стационарной модели невозможным. Причиной этой проблемы является то, что в формуле (1) присутствуют отсчеты при $k > 0$, которые приходится исключать по мере приближения к отсчету t .

Предлагаемым способом устранения этого недостатка является использование односторонних фильтров, т. е. таких фильтров, которые подразумевают суммирование только по предшествующим текущему отсчетам. Проблемы с суммированием остаются при $t < n$, однако при достаточно большом T первые n отсчетов в аппроксимирующей и детализирующей компонентах можно отбросить без ущерба для точности прогноза.

Применение односторонних фильтров обосновано и с методологической точки зрения – построенный с их помощью прогноз не зависит от будущего. В непрерывном случае односторонним фильтрам соответствуют финитные функции φ , область компактного задания которых лежит на отрицательной полуоси. Иными словами, в формулах (3), (4) интегрирование ведется от $-\infty$ до 0.

Оценка полиномиального тренда. Пусть случайный процесс $c(t)$ имеет вид $m(t) + \xi_t$, где $m(t) = \sum_{i=0}^s a_i t^i$ – полиномиальный тренд, а ξ_t – стационарный случайный процесс. В этих предположениях справедлива следующая

Теорема. Существует функция $\varphi(v)$, компактным носителем которой является интервал $[-1, 0]$, такая, что выполняется условие нормировки (2), и математическое ожидание детализирующего процесса тождественно равно нулю. При этом такая функция может быть найдена в виде

$$\varphi(v) = \begin{cases} v^{s+1} - \sum_{i=0}^s q_i v^i, & v \in [-1, 0], \\ 0, & v \notin [-1, 0], \end{cases} \quad (6)$$

где коэффициенты q_i находятся из системы

$$\begin{cases} \alpha_0 q_0 + \alpha_1 q_1 + \dots + \alpha_s q_s = \alpha_{s+1} - 1 \\ \alpha_1 q_0 + \alpha_2 q_1 + \dots + \alpha_{s+1} q_s = \alpha_{s+2} \\ \alpha_2 q_0 + \alpha_3 q_1 + \dots + \alpha_{s+2} q_s = \alpha_{s+3} \\ \vdots \\ \alpha_s q_0 + \alpha_{s+1} q_1 + \dots + \alpha_{2s} q_s = \alpha_{2s+1} \end{cases}, \quad \alpha_i = \frac{(-1)^i}{i+1}, i = 0, 1, 2, \dots, \quad (7)$$

которая имеет единственное решение.

Доказательство. Математическое ожидание детализирующего процесса имеет вид [6]:

$$n(t) = \sum_{i=0}^s a_i \left(t^i - \sum_{j=0}^i C_i^j t^{i-j} \int_R v^j h(v) dv \right).$$

Поскольку функция $n(t)$ должна быть константой при всех возможных значениях вектора коэффициентов a_i , необходимо и достаточно, чтобы все интегралы $\int_{-\infty}^0 v^j h(v) dv$ обращались в 0 для всех $j = 1, \dots, s$. Это равносильно тому, что

$$\int_{-\infty}^0 v^j \varphi(v) dv = 0, j = 1, \dots, s. \quad (8)$$

Подставляя выражение (6) в условия (2), (8), получаем систему

$$\begin{aligned} \int_{-1}^0 \left(v^{s+1} - \sum_{i=0}^s q_i v^i \right) dv &= 1; \\ \int_{-1}^0 v^j \left(v^{s+1} - \sum_{i=0}^s q_i v^i \right) dv &= 0, j = 1, \dots, s. \end{aligned} \quad (9)$$

Обозначим $\alpha_i = \int_{-1}^0 v^i dv = \frac{(-1)^i}{i+1}$, $i = 0, 1, 2, \dots$. Тогда после раскрытия скобок система (9) принимает

вид (7). Разрешив эту систему относительно q_i , можно построить функцию φ в соответствии с формулой (6). Покажем разрешимость системы (7) при любом s . Заметим, что все главные миноры матрицы системы являются определителями аналогичных матриц меньших порядков. Таким образом, достаточно показать, что матрица системы положительно определена. Рассмотрим соответствующую квадратичную форму

$$Q = \sum_{i=1}^{s+1} \sum_{j=1}^{s+1} \alpha_{i+j-2} x_i x_j = \sum_{i=1}^{s+1} \sum_{j=1}^{s+1} \frac{(-1)^{i+j}}{i+j} x_i x_j.$$

Воспользуемся соотношением $\frac{1}{a} = \int_0^{\infty} e^{-at} dt$, $a > 0$. Тогда

$$Q = \sum_{i=1}^{s+1} \sum_{j=1}^{s+1} \int_0^{\infty} (-1)^{i+j} e^{-(i+j)t} dt x_i x_j = \int_0^{\infty} \left(\sum_{i=1}^{s+1} (-1)^i e^{-it} x_i \right)^2 dt.$$

Подынтегральная функция неотрицательна, поэтому весь интеграл также неотрицателен, что доказывает неотрицательную определенность формы. Если же интеграл равен нулю, в силу неотрицательности и непрерывности подынтегральная функция также должна быть тождественно равна нулю. Поэтому равенство нулю всей формы равносильно тому, что $\sum_{i=1}^k (-1)^i e^{-it} x_i = 0$ при любом t , но, поскольку функции e^{-it} линейно независимы, последнее возможно лишь при всех x_i , равных 0, что и завершает доказательство положительной определенности квадратичной формы.

Замечание. Очевидно, функция, удовлетворяющая всем условиям (8) для некоторого $s = s_0$, будет удовлетворять им и при всех $s < s_0$.

Пример. Решая систему (9) при $s = 3$, получаем фильтр φ , выделяющий полиномиальный тренд степени не выше $s = 3$, вида

$$\varphi(t) = \begin{cases} t^4 + 142t^3 + 1689/7t^2 + 842/7t + 1121/70, & t \in [-1, 0], \\ 0, & t \notin [-1, 0]. \end{cases} \quad (10)$$

По данной функции можно построить односторонний фильтр $\phi(k) = \varphi(k/n) / \sum_{k=-n}^0 \varphi(k/n)$, $k = -n, -n+1, \dots, 0$ (здесь значения фильтра при $k > 0$ не указаны, так как равны 0). При $n = 5$, например, получаем фильтрующий вектор, равный $(-0,331, 0,159, 0,014, -0,213, 0,040, 1,332)$.

Одноуровневый алгоритм прогнозирования и вычислительный эксперимент. Рассмотрим модель случайного процесса $c(t) = m(t) + \xi_t$, $t \in N$, где ξ_t – стационарный шум. Приведем упрощенный вариант алгоритма многоуровневого прогнозирования – одноуровневый алгоритм прогнозирования временных рядов с использованием непрореженного вейвлет-преобразования.

Пусть $Y(t)$ – наблюдения процесса $c(t)$ в моменты времени $t = 1, \dots, T$ и требуется построить прогноз $\hat{Y}(t)$ значений $c(t)$ в моменты времени $t = T+1, \dots, T+\tau$, $\tau \in N$. С помощью дискретного непрореженного вейвлет-преобразования, используя фильтр-вектор, вычисляемый на основании фильтрующей функции φ по формуле $\phi_{-k} = \varphi(-k/n)$, $k = 0, 1, \dots, n$, построим аппроксимирующую и детализующую компоненты – ряды $D(t)$ и $W(t)$.

Компоненту $D(t)$ примем за оценку тренда $m(t)$ и будем экстраполировать ее, оценивая коэффициенты полинома с помощью метода наименьших квадратов (МНК). Компоненту $W(t) = Y(t) - D(t)$ примем за оценку шума и будем прогнозировать ее с использованием модели АР. Например, если порядок авторегрессии известен, для оценки коэффициентов АР можно использовать метод Юла – Уокера. Окончательная оценка будет иметь вид $\hat{Y}(t) = \hat{D}(t) + \hat{W}(t)$, $t = T+1, \dots, T+\tau$, где $\hat{D}(t) = \sum_{i=0}^s \hat{a}_i t^i$ – оценка тренда, в которой коэффициенты $\hat{a}_i$ построены по значениям $D(t)$ при $t = 1, \dots, T$, а $\hat{W}(t)$ – упомянутый прогноз $W(t)$.

При проведении эксперимента моделировался временной ряд $Y(t) = m(t) + \xi_t$, $t = 1, \dots, T+\tau$, где $m(t) = t^3 - t^2 + 2t + 3$, $\xi_t = \gamma \zeta_t$, ζ_t – процесс АР(q) с коэффициентами $b_1, \dots, b_q$ и инновационным процессом со стандартным гауссовским распределением; γ – коэффициент, регулирующий дисперсию шума. Затем по первым T значениям ряда $Y(t)$ происходило прогнозирование следующих τ отсчетов. Для прогнозирования тренда использовались оценки коэффициентов полинома по методу МНК при известной степени полинома, для коэффициентов АР использовались оценки Юла – Уокера при известном порядке авторегрессии. Для каждого значения γ проводилось $N = 1000$ экспериментов.

В каждом эксперименте описанным способом генерировалась выборка, к которой затем применялись четыре способа прогнозирования:

- 1) одноуровневое прогнозирование с использованием обычного фильтра Хаара ($\varphi(t) = 0,5I_{[-1,1]}(t)$), как в работе [1];
- 2) одноуровневое прогнозирование с использованием одностороннего фильтра Хаара ($\varphi(t) = I_{[-1,0]}(t)$);
- 3) одноуровневое прогнозирование с использованием фильтра (9) из [6];
- 4) одноуровневое прогнозирование с использованием фильтра (10).

В 1) и 3) приведены данные из [6] для сравнения результатов.

Для оценки качества прогноза вычислялось среднее относительное отклонение

$$\delta = \frac{1}{\tau} \sum_{t=T+1}^{T+\tau} |Y(t) - \hat{Y}(t)|.$$

Для каждого из методов прогнозирования это отклонение усреднялось по N экспериментам. Использовались следующие модельные значения: $T = 100$, $\tau = 2$, $q = 2$, $b_1 = -0,25$, $b_2 = 1$. Значение параметра n для построения дискретного фильтра было взято равным 5. Результаты эксперимента приведены в таблице.

Результаты численного эксперимента

γ	1	100	1000	3000	6000
δ (фильтр Хаара [1])	0,0670	0,0670	0,0673	0,0679	0,0674
δ (односторонний фильтр Хаара)	0,0311	0,0312	0,0312	0,0322	0,0339
δ (фильтр (9) из [6])	0,0300	0,0300	0,0295	0,0314	0,0500
δ (фильтр (10))	0,0176	0,0176	0,0177	0,0226	0,0346

Как следует из результатов, использование фильтра (10) обеспечивает наибольшую точность прогноза.

* * *

Приведен способ устранения краевых эффектов при НПВП с помощью одностороннего фильтра. Построен пример одностороннего фильтра для частного случая полиномиального тренда и показано, что использование таких фильтров существенно повышает точность прогноза по сравнению с двусторонними фильтрами в сочетании с методом уменьшения базы. На примере показано, что при применении непрореженного вейвлет-преобразования правильный выбор фильтра существенно влияет на точность прогноза. Доказано, что в случае модели с полиномиальным трендом и стационарным шумом удовлетворяющий критерию фильтр всегда может быть построен и представлен алгоритм построения такого фильтра.

1. Renaud O., Starck J.-L., Murtagh F. // Cahiers du department e'conometrie Faculte des sciences economiques et sociales Universite de Geneve. 2004. № 4. http://www.unige.ch/ses/metri/cahiers/2002_04.pdf. Дата доступа: 20.10.2010.

2. Conejo A.J., Plazas M.A., Espinola R., Molina A.B. // IEEE Transactions on Power Systems. 2005. Vol. 20. № 2. P. 1035.

3. Fryzlewicz P., Van Belleghem S., Von Sachs R. // Annals of the Institute of Statistical Mathematics. 2003. № 55. P. 737.

4. Aussem A., Campbell J., Murtagh F. // J. of Computational Intelligence in France. 1998. № 6 (2). P. 5.

5. Craigmile P.F., Percival D.B. // Encyclopedia of Environmetrics / Ed. by A.H. El-Shaarawi, W.W. Piegorsch. Chichester, 2002. Vol. 4.

6. Миротин Е.А. // Информатика. 2010. № 4 (28). С. 48.

Поступила в редакцию 23.03.11.

Евгений Адольфович Миротин – аспирант кафедры математического моделирования и анализа данных. Научный руководитель – кандидат физико-математических наук, доцент, заведующий НИЛ статистического анализа и моделирования НИИ прикладных проблем математики и информатики БГУ М.С. Абрамович.