

УДК: 681.3.069

Т.Г. Лазебник

ПРОГНОЗИРОВАНИЕ ДИНАМИЧЕСКИХ РЯДОВ С СЕЗОННОЙ КОМПОНЕНТОЙ С ИСПОЛЬЗОВАНИЕМ ВСТРОЕННОЙ ФУНКЦИИ EXCEL АНАЛИЗ ФУРЬЕ

У цій статті розглядаються питання використання перетворення Фур'є для виділення частотних складових спектру сигналу з метою його подальшого розділення та використання для прогнозування економічних процесів із вираженою сезонною компонентою.

This article examines the issues of Furier transformation usage in order to pick out frequency components of signal spectrum with the aim of its further separation and usage for forecasting economic processes with expressive seasonal component.

Ключевые слова: преобразование Фурье, частотная область представления, сезонность, прогнозирование.

Известно, что при прогнозировании процессов, подверженных сезонным колебаниям, возникают определенные трудности. Например, методы регрессионного анализа, которые дают хорошие результаты для стационарных и квазистационарных процессов, здесь работают плохо. При исследовании сезонных колебаний для достижения желаемых результатов стараются применять другие методы, связанные, например, с вычислениями индексов сезонности или иных коэффициентов [1; 2; 3]. Некоторые из используемых методов довольно сложные, например метод Бокса — Дженкинса.

Здесь предлагается рассмотреть применение функции Excel *Анализ Фурье* для прогнозирования динамических рядов с выраженной сезонной составляющей. Возможность использования данной функции основана на следующем.

Временные ряды статистических наблюдений, выполняемых через равные периоды времени, можно рассматривать как дискретный временной сигнал $\{x_0, x_1, x_2, \dots, x_{N-1}\}$.

Известно, что любой дискретный сигнал может быть представлен в виде суммы гармоник, имеющих разные частоты и амплитуды [4]. При этом количество гармонических составляющих будет конечно, так как интервал наблюдений ограничен во времени, и равно $N/2$, где N — число наблюдений во временном ряду.

Основным свойством преобразования Фурье [5; 6] является тот факт, что из последовательности значений $\{x_0, x_1, x_2, \dots, x_{N-1}\}$, расположенных во временной области, при прямом преобразовании Фурье

$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi kn/N} \quad (1)$$

получается последовательность спектральных составляющих $\{X_0, X_1, X_2, \dots, X_{N-1}\}$, соответствующих параметрам каждой гармоники. И, таким образом, исходный временной сигнал переводится в частотную область.

Если затем к полученной последовательности $\{X_0, X_1, X_2, \dots, X_{N-1}\}$ применить обратное преобразование Фурье

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{j2\pi kn/N}, \quad (2)$$

то сигнал снова будет переведен во временную область. При этом сумма гармоник будет совпадать с исходной последовательностью $\{x_0, x_1, x_2, \dots, x_{N-1}\}$ временного ряда [5; 6].

Здесь предлагается после перехода в частотную область с помощью фильтрующего преобразования разделить спектр сигнала на спектр с низшими гармониками и на спектр с высшими. А затем с помощью обратного преобразования Фурье возвратиться к временному представлению, но уже не одного, а двух сигналов — низкочастотного и высокочастотного. Сумма этих сигналов также будет совпадать с исходной последовательностью временного ряда. Следует ожидать, что каждая из полученных составляющих будет более гладкой по сравнению с исходным сигналом. А поэтому к каждой из них для прогнозирования можно будет применять методы, например, регрессионного анализа с ожиданием более адекватного результата. После проделанных процедур обе составляющие необходимо объединить и получить, таким образом, итоговую кривую с прогнозом.

Технологию выполнения изложенного алгоритма рассмотрим на конкретном примере. В качестве исходных данных был выбран динамический ряд, представляющий сезонные колебания цен на лекарственный препарат пинсол в каплях, взятый на сайте [7]. Данные приведены в табл. 1. Построенная по ним базовая кривая показана на рис. 1. Этот препарат относится к группе лекарств, обладающих ярко выраженным сезонным характером ценовых изменений.

Таблиця 1

Фактические средние розничные цены на «Пиносол» в каплях по 10 мл в 60 аптеках Санкт-Петербурга, 2000–2002 гг.

(источник: <http://www.dis.ru/library/market/archive/2006/3/4227.htm>)

месяц	Средние розничные цены, руб.		
	2000 г.	2001 г.	2002 г.
январь	20,65	22,54	26,67
февраль	20,16	22,12	25,74
март	20,47	21,99	25,56
апрель	20,09	21,8	25,24
май	19,97	21,62	24,93
июнь	20,23	21,7	25,56
июль	19,87	21,53	24,87
август	20,91	22,22	25,45
сентябрь	21,17	22,52	25,97
октябрь	23,3	23,52	26,44
ноябрь	23,01	23,19	25,79
декабрь	22,48	23,37	25,14

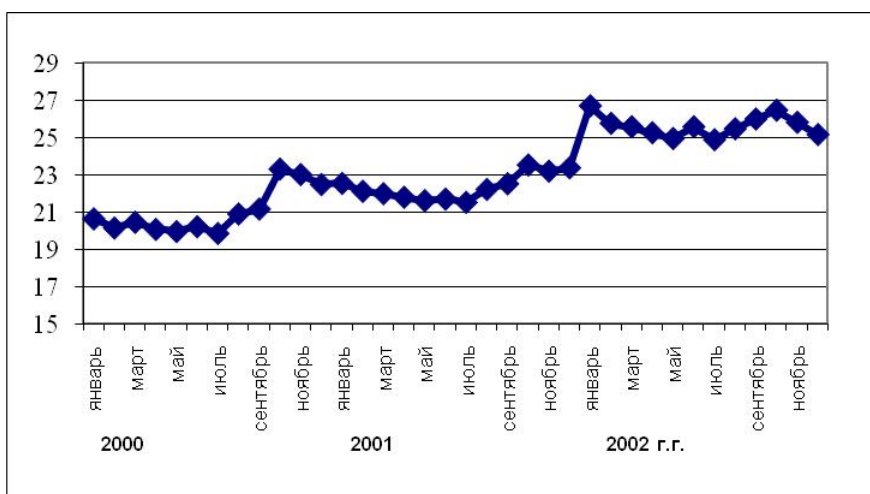


Рис. 1. Фактические средние розничные цены на «Пиносол» в каплях за период 2000–2002 г.г. (разработано автором на основе источника [7])

Согласно предложенной методике, для прогнозирования ценовых изменений необходимо представить базовую кривую как совокупность низкочастотного сигнала, обусловленного влиянием сезонности, и высокочастотного, обусловленного присутствием остальных гармоник. Таким образом, кривую, показанную на рис. 1, необходимо разложить на две составляющие — низкочастотную и высокочастотную. Выполнить это возможно с использованием функции *Анализ Фурье*, который позволит

получить спектр рассматриваемого временного ряда и определить из каких гармонических составляющих состоит рассматриваемая базовая кривая.

Функция *Анализ Фурье* работает с массивами данных и реализует алгоритм быстрого преобразования Фурье (БПФ). Обязательным условием работы алгоритма БПФ является то, что число членов N исходного массива должно быть равно 2^n , где $n = 1, 2, 3, \dots$. Количество исходных точек в нашем примере согласно табл. 1 — $12 \times 3 = 36$. Поскольку цель данной статьи показать как работает предлагаемый метод и оценить качество полученного прогноза, процесс будем рассматривать с января 2000 г. по август 2002 г. включительно. На период с сентября по декабрь 2002 г. сделаем прогноз и сравним полученный результат с фактическими данными, приведенными в табл. 1.

Разложим базовую кривую на гармонические составляющие. Для этого обратимся к функции *Анализ Фурье*, которая позволяет получить комплексный спектр сигнала и осуществить, таким образом, переход от временного представления базовой кривой к ее спектральному виду. На рис. 2 из-за ограниченного объема статьи показаны только начальные значения исходных данных и полученного комплексного спектра. Для наглядности можно вычислить модуль спектра, воспользовавшись функцией *Excel/МНИМ.ABS()*. Результат показан в виде графика на рис. 3, из которого видно, что спектр базовой кривой состоит из 16 гармонических составляющих разной частоты.

Получив вещественную и мнимую части спектра базового сигнала, можно приступить к его разделению на высокочастотную и низкочастотную компоненты.

	A	B	C	D	E	F	G
4							
5	год	месяц	Средние разничные цены, руб.	Комплексный спектр сигнала			
6	2000	январь	20,65	724,45			
7		февраль	20,16	3,16398552508603+29,8673370760104i			
8		март	20,47	-11,9845900956596+24,8259689930656i			
9		апрель	20,09	2,58076833531937+4,83931997830576i			
10		май	19,97	0,835893577750956+0,7854267658087i			
11		июнь	20,23	-2,49733679051978+5,5911496878864i			
12		июль	19,87	-3,85506975825358+8,15581131760411i			
13		август	20,91	-3,20715120595319+0,802085276955311i			
14		сентябрь	21,17	-0,42000000000016-0,7700000000001i			
15		октябрь	23,3	-1,73271197652534+4,92640911635899i			
16		ноябрь	23,01	-3,36755867397294+3,45041667579513i			
17		декабрь	22,48	-2,03213121337683-1,2101025342133i			
18	январь	22,54	-0,535893577750947+2,08547676580868i				

Рис. 2. Базовый сигнал и его комплексный спектр (разработано автором на основе источника [7])

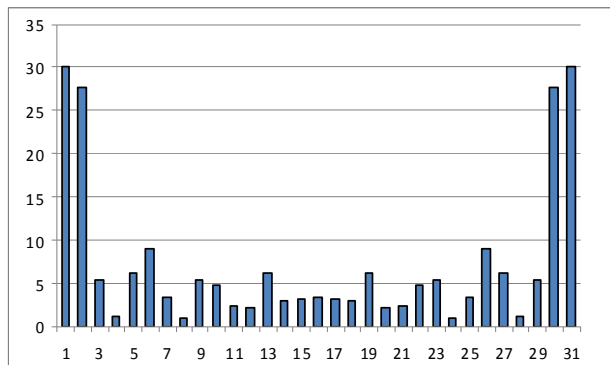
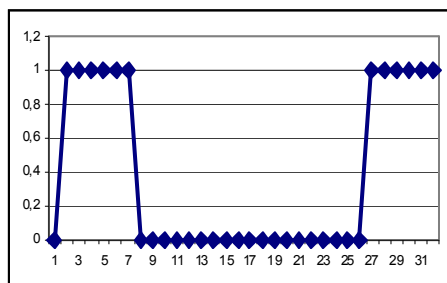
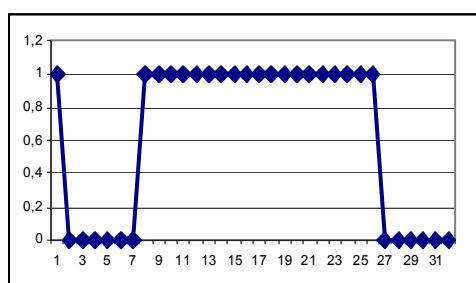


Рис. 3. Модуль комплексного спектра (розроблено автором)

Ета процедура хорошо известна и широко используется в области цифровой фильтрации [8]. В нашем случае необходимо создать 2 фильтра — один для выделения низкочастотной составляющей сигнала, другой — для высокочастотной. На рис. 4 показаны используемые фильтры. Из рисунка видно, что для формирования низкочастотного сигнала будут использованы первые 6 низкочастотных гармоник, начиная со второй, а для высокочастотного — все остальные. Необходимо отметить, что не существует четкого критерия, по которому бы выбиралось количество гармоник для каждой из составляющих. Решение здесь принимается исследователем, исходя из характера процесса.



а) НЧ фильтр



б) ВЧ фильтр

Рис. 4. Фильтры, используемые для разделения базового сигнала на низкочастотную (а) и высокочастотную (б) составляющие (разроблено автором)

После выполнения комплексного умножения каждого из фильтров на комплексный спектр базовой кривой с помощью функции `МНИМ.ПРОИЗВЕД()`, сформируются комплексные спектры двух составляющих базовой кривой — низкочастотной и высокочастотной. Чтобы перейти от спектрального представления полученных компонент к их представлению во временной области, необходимо выполнить обратное преобразование Фурье [5]. Для этого в диалоговом окне функции *Анализ Фурье* устанавливается флажок *Инверсия* и выполняется процедура вычисления БПФ.

На рис. 5 приведені розраховані речові складові даних сигналів, отримані за допомогою функції МНІМ.ВЕЩ(), і загальний сигнал, що включає адитивні складові — низькочастотну і високочастотну компоненти: ІТОГОВИЙ СИГНАЛ = НЧ сигнал + ВЧ сигнал

Графічні зображення низькочастотної і високочастотної складових показані на рис. 6. Розмах варіацій по амплітуді між максимумом і мінімумом в низькочастотному сигналі становить $\sim 6,5$ одиниць і при цьому спостерігається його періодичність. А розмах варіацій для високочастотного сигналу становить всього лише $\sim 2,5$. Характер сигналу при цьому явно шумовий. Якщо порівняти значення загального сигналу з базовим на інтервалі спостереження, то можна побачити, що вони збігаються. Наша ж задача полягає в отриманні прогнозних значень за межами інтервалу спостереження.

Речові частини НЧ сигналу	Речові частини ВЧ сигналу	Загальний сигнал
-0,734773	21,384772	20,65
-2,428131	22,58813	20,16
-2,878456	23,348456	20,47
-2,490009	22,580008	20,09
-2,244722	22,214721	19,97
-2,523803	22,753803	20,23
-2,741923	22,611924	19,87
-2,163149	23,07315	20,91
-0,871474	22,041476	21,17
0,2584877	23,041514	23,3
0,5101991	22,499803	23,01

Рис. 5. Низькочастотна і високочастотна складові сигналу в часовій області, а також загальний сигнал (розроблено автором)

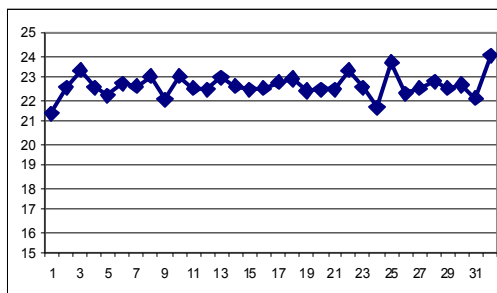
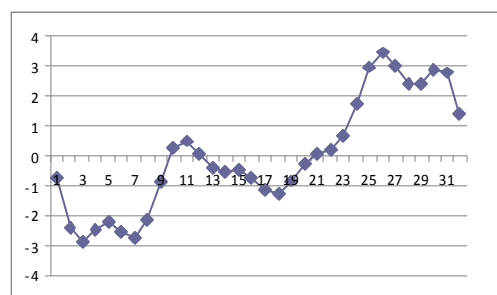


Рис. 6. Низькочастотна складові базового сигналу (вгорі) і високочастотна складові базового сигналу (вниз) (розроблено автором)

Прогнозування шумового процесу з невеликим розмахом варіацій можна виконати за допомогою методів регресійного аналізу. Тут цілком доцільно скористатися можливостями Excel. Для цього на графіку високочастотного сигналу за допомогою опції *Додати лінію тренда* побудуємо лінійну апроксимуючу криву з вказанням прогнозу на 4 періода вперед і з вимогами показати рівняння на діаграмі. Нагадаємо, що 4 періода вибрані тому, що ми маємо реальні дані на період вересень—грудень 2002 г., які можна порівняти з

результатом прогноза. На рис. 7 показана полученная линейная аппроксимация высокочастотного сигнала.

Поскольку у низкочастотного сигнала большая вариабельность и выраженная периодичность, то применять для прогноза методы регрессионного анализа с использованием всей базовой кривой нецелесообразно, так как может быть получен трудно интерпретируемый результат. С учетом того, что сигнал меняется довольно плавно, предлагается использовать только последние несколько точек низкочастотного сигнала и по ним выполнить прогноз на последующие периоды. Сразу отметим, что нет четкого критерия для определения количества необходимых точек для прогноза. Выбор целесообразно производить с учетом общих правил, применяемых при прогнозировании. В данном случае были выбраны 8 последних точек низкочастотной кривой и с помощью опции *Добавить линию тренда* выполнена аппроксимация с помощью степенной функции с прогнозом на 4 периода вперед. На рис. 8 приведен результат выполненных действий.

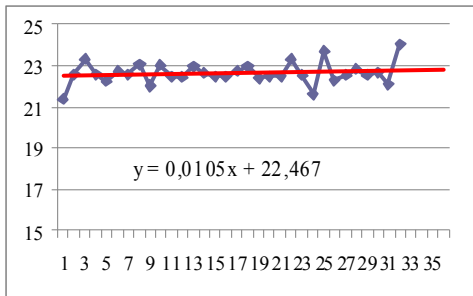


Рис. 7. Высокочастотная составляющая базового сигнала и его линейная аппроксимация с прогнозом на 4 последующие периода (разработано автором)

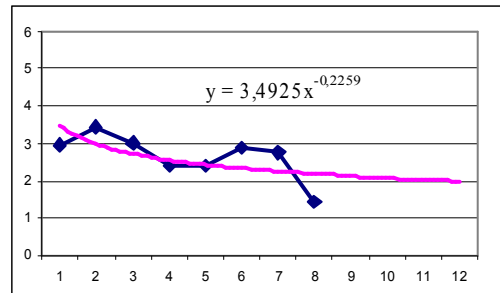


Рис. 8. Последние 8 точек низкочастотной составляющей базового сигнала и их аппроксимация с помощью степенной функции с прогнозом на 4 последующие периода (разработано автором)

Уравнения аппроксимации:

$$Y_{\text{вч}} = 0,0105x + 22,467$$

$$Y_{\text{нч}} = 3,4925x(-0,2259)$$

— для высокочастотной составляющей;

— для низкочастотной составляющей;

По данным уравнениям рассчитываются значения 4-х прогнозируемых периодов для каждой составляющей и вычисляется сумма этих значений. На рис. 9 представлен полученный таким образом итоговый сигнал с прогнозом на сентябрь–декабрь 2002 г.

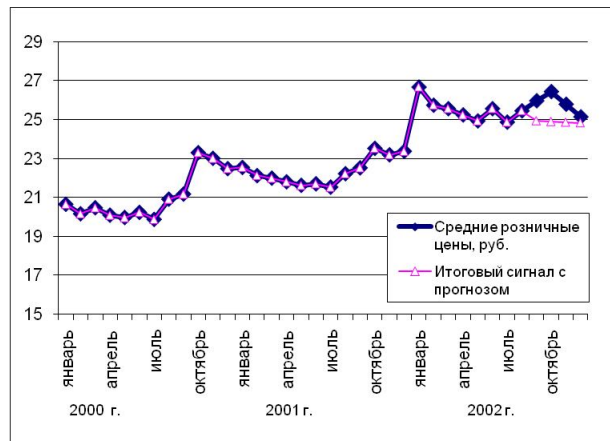


Рис. 9. Базовый сигнал, представляющий розничные цены, и итоговый сигнал с прогнозом, полученный в результате моделирования с использованием функции *Анализ Фурье* (разработано автором)

В таблице 2 приведены числовые значения, рассчитанные по данным уравнениям, а также рассчитана ошибка прогноза для этих 4-х точек. Определим среднюю точность аппроксимации построенной модели. Из табл. 1 и рис. 5 и 9 видно, что значения фактических цен и полученных по модели до сентября 2002 г. совпадают, поэтому говорить о точности на данном интервале излишне. И только для прогнозных значений с сентября до декабря 2002 г. наблюдаются расхождения между фактическими значениями и прогнозом, которые можно оценить.

Таблица 2

Прогнозные значения уровня цен и ошибка прогноза (получено автором в результате обработки сигнала)

Период.	НЧ составляющая	ВЧ составляющая	Уровни цен, полученные по модели $? = Y_{НЧ} + Y_{ВЧ}$	Фактические цены, руб.	Ошибка прогноза, % $((Y - ?) / Y) 100\%$
2002 г.	$Y_{НЧ}$	$Y_{ВЧ}$		Y	
сентябрь	2,126048	22,8135	24,94	25,97	3,97 %
октябрь	2,076043	22,824	24,90	26,44	5,82 %
ноябрь	2,031823	22,8345	24,87	25,79	3,58 %
декабрь	1,992275	22,845	24,84	25,14	1,20 %

Средняя точность аппроксимации, определяемая по формуле [3]:

$$A = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| 100\%$$

для данного интервала равна 3,64 %.

Предлагаемая методика довольно проста и доступна. Ее можно использовать для разработки краткосрочных прогнозов с высокой точностью целого ряда экономических показателей, характеризующихся сезонностью изменений. При этом сочетание знаний экономической природы исследуемых процессов и творческого подхода к применению предлагаемой методики позволит получать надежные и достоверные прогнозы.

1. *Карлберг К.* Бизнес-анализ с помощью Microsoft Excel / К. Карлберг; пер. с англ. — М.: Изд. дом «Вильямс», 2005. — 464 с.; 2. *Гусаров В.М.* Статистика: Учебное пособие для ВУЗов / В.М. Гусаров. — М.: ЮНИТИ — ДАНА, 2003. — 463 с.; 3. *Дуброва Т.А.* Статистические методы прогнозирования в экономике: Учебное пособие, практикум, тесты, программа курса / Т.А. Дуброва, М.Ю. Архипова. — М.: Московский государственный университет экономики, статистики и информатики, 2004. — 136 с.; 4. *Френкс Л.* Теория сигналов / Л. Френкс. — М.: Советское радио, 1974. — 344 с.; 5. *Айфичер Э.С.* Цифровая обработка сигналов: практический подход / Э.С. Айфичер, Б.У. Джервис; пер. с англ. — [2-е изд.]. — М.: Издательский дом «Вильямс», 2004. — 992 с.; 6. *Марпл-мл. С.Л.* Цифровой спектральный анализ и его приложения / С.Л. Марпл-мл.; пер. с англ. — М: Мир, 1990. — 584 с.; 7. *Орлов А.С.* Прогнозирование сезонных колебаний цен на лекарственные средства: [Электронный ресурс]. — Режим доступа: <http://www.dis.ru/library/market/archive/2006/3/4227.htm>; 8. *Отнес Р.* Прикладной анализ временных рядов. Основные методы / Р. Отнес, Л. Эноксон; пер. с англ. — М.: Мир, 1982. — 428 с.

УДК: 332.1(075.8)+51(075.8)