

© Алексей Николаевич Кисляков
канд. техн. наук, доцент,
Владимирский филиал РАНХиГС

ВЛИЯНИЕ ФИЛЬТРАЦИИ ФУРЬЕ НА ТОЧНОСТЬ ПРОГНОЗА СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Аннотация: рассматривается возможность удаления шумовых компонент временных рядов социально-экономических показателей с использованием дискретного преобразования Фурье и применением оконной фильтрации. Проведено исследование эффективности метода для сгенерированных временных рядов с аддитивной шумовой составляющей и без нее. Качество фильтрации временного ряда определяется видом окна фильтрации, от которого также зависит и точность дальнейшего прогноза значений временного ряда.

Abstract: Discusses the possibility of removing the noise component of the time-series social and economic indicators, using the discrete Fourier transform and applying a window filter. A study of the effectiveness of the method for the generated time series with the additive noise component and without it. The quality of filtering of the time series is determined by the type of the filtering window, which also depends on the accuracy of future forecast values of a time series.

Ключевые слова: быстрое преобразование Фурье, временные ряды, дискретная фильтрация.

Keywords: fast Fourier transform, time series, discrete filtering.

На сегодняшний день насчитывается множество подходов к анализу и моделированию временных рядов, являющихся важнейшими источниками информации о закономерностях развития различных социально-экономических процессов.

Наиболее популярными являются методы факторного анализа, требующие априорной информации о характере исследуемого временного ряда [1]. Известно, что любая социально-экономическая система является кибернетической системой открытого типа, поэтому на исследуемый процесс оказывает влияние постоянно меняющийся набор факторов, вследствие чего во временных рядах присутствуют различного рода шумы. Шумовая составляющая может быть представлена в виде аддитивной и мультипликативной модели. Существуют методы выявления отдельных аномальных уровней временных рядов и их сглаживания, однако эти методы не позволяют компенсировать шумы на протяжении всего интервала наблюдения.

Кроме того, в данном случае неизбежно возникает проблема получения точного прогноза ввиду того, что спрогнозированные данные также будут сглажены.

Поэтому наиболее подходящим методом удаления шумовой компоненты является Фурье-фильтрация данных, позволяющая выполнить сглаживание временного ряда путем удаления высокочастотных спектральных компонент временного ряда, которые характеризуют шумы различных типов.

Преобразование Фурье связывает временной или пространственный сигнал (или же некоторую модель этого сигнала) с его представлением в частотной области [2].

Так как отсчеты временных рядов обычно фиксируются через определенные интервалы времени, то временной ряд называется дискретным, и описанное преобразование Фурье, примененное к дискретной последовательности сигналов, называется дискретным преобразованием Фурье (ДПФ). Так для последовательности из N точек спектр ДПФ вычисляется по формуле:

$$X[k] = \frac{1}{N} \cdot \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi \frac{nk}{N}}, \quad (1)$$

где n – дискретное время сигнала, k – отсчеты по частоте сигнала, $x(n)$ – исследуемая временная последовательность, $X(k)$ – ее спектр. Выражение

$$W_N = e^{-j2\pi \frac{nk}{N}} \quad (2)$$

называется комплексным оператором ДПФ. Для вычисления одного k -го отсчета используется все N временных отсчетов $x(n)$, для чего требуется N^2 процедур комплексных операций сложения и умножения.

Для быстрого вычисления спектра дискретно-временного ряда используется быстрое преобразование Фурье (БПФ), которое заключается в разделении элементов последовательности $x(n)$ на четные и нечетные компоненты и получение 2 точечных ДПФ четных и нечетных отсчетов.

$$\begin{aligned} X[k] &= \frac{1}{N} \cdot \sum_{m=0}^{N/2-1} x[2m] \cdot W_N^{2mk} + \frac{1}{N} \cdot W_N^k \cdot \sum_{m=0}^{N/2-1} x[2m+1] \cdot W_N^{2mk} = \\ &= X_{N/2}[k, 0] + W_N^k \cdot X_{N/2}[k, 1], \end{aligned} \quad (3)$$

где каждое из слагаемых является преобразованием длины $N/2$.

$$X_{N/2}[k, p] = \frac{1}{N} \cdot \sum_{n=0}^{N/2-1} x[2n+p] \cdot W_{N/2}^{2k}, \quad p=0,1. \quad (4)$$

Возможно продолжить дальнейшее прореживание по времени до достижения значения количества отсчетов $N/2 = 2$, тогда если на каждом этапе выполняется $N/2$ операций умножения и сложения, то количество операций в алгоритме БПФ будет равно $(N/2) \cdot \log_2 N$.

Методика прогнозирования социально-экономических показателей на основе БПФ предполагает выявление детерминированной компоненты временных рядов [3] путем фильтрации шумовых составляющих и включает следующие шаги.

1. Посредством БПФ вычисляется дискретный спектр временного ряда с целью выделения высокочастотных компонент.

2. Выбирается тип и параметры «скользящего» окна фильтрации.

3. С помощью фильтра нижних частот удаляется составляющая, соответствующая шумовым компонентам сигнала.

4. Выполняется обратное преобразование Фурье и восстанавливается исходная временная последовательность для последующего построения прогноза.

5. Вычисляется ошибка восстановления. Сравнивается с заданной. Корректируется модель фильтра.

Рассмотрим пример реализации с помощью пакета MATLAB алгоритма прогнозирования стационарных временных рядов с применением преобразования Фурье.

На начальном этапе задается модель временного ряда [4], состоящего из нескольких гармоник (циклических компонент) с разными частотами и аддитивной случайной компоненты с нормальным законом распределения:

$$x(t) = a_1 \cos(\omega_1 \cdot t) + a_2 \cos(\omega_2 \cdot t) + \dots + a_n \sin(\omega_n \cdot t) + c \cdot \text{randn}(t). \quad (5)$$

Значения A_i , ω_i и C выбираются произвольно. Так, одна из реализаций исходного сигнала имеет следующий вид:

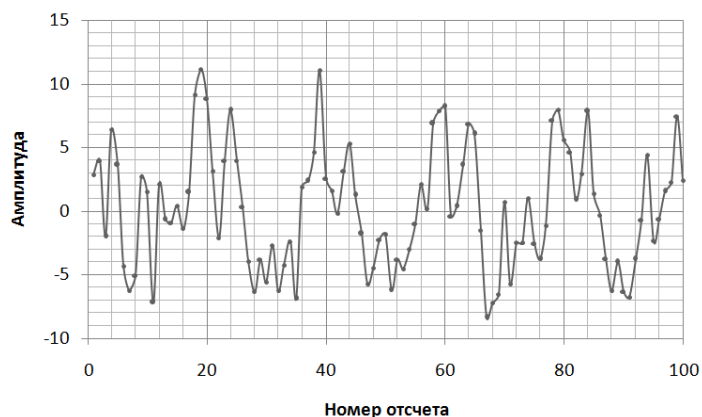
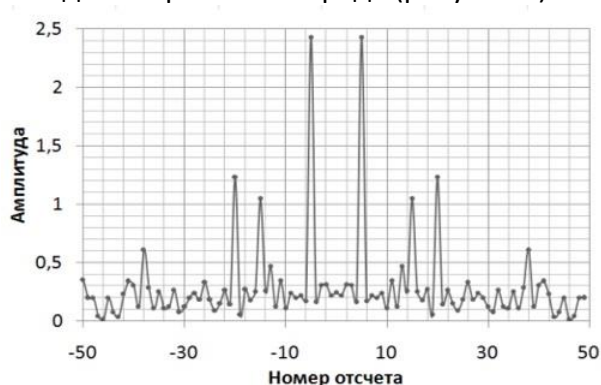
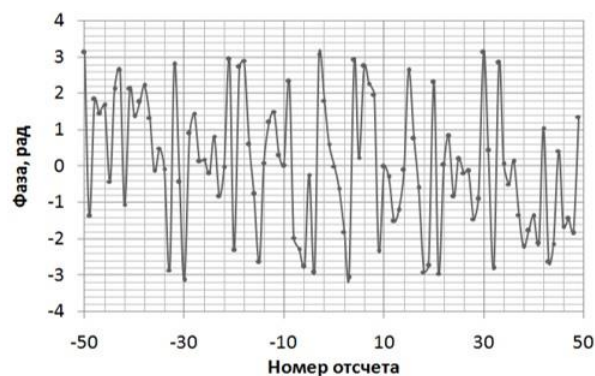


Рис. 1. Одна из вариаций исходного временного ряда

На следующем этапе выполняется ДПФ и строится амплитудный и фазовый спектры исходного временного ряда (рисунок 2).



а)



б)

Рис. 2. Спектр временного ряда: а) – амплитудный; б) фазовый

На полученной диаграмме амплитудного спектра отчетливо прослеживаются те гармонические составляющие ряда, которые имеют наибольшую амплитуду. Без труда можно определить и их частоты и восстановить значения временного ряда, суммируя эти гармоники, получив сигнал $y(t)$.

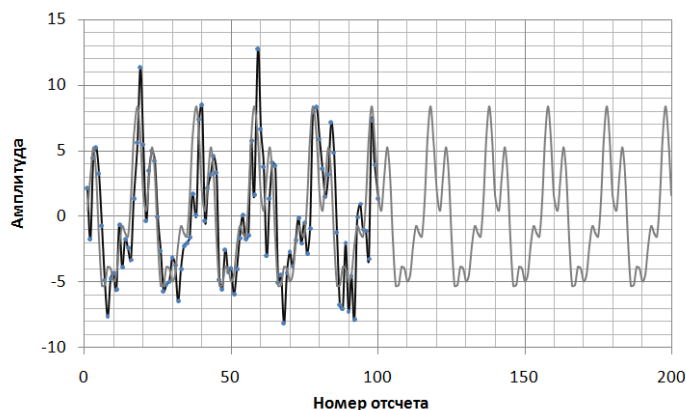


Рис. 3. Исходное и прогнозируемое на 100 отсчетов вперед представление временного ряда

Полученный временной ряд имеет сезонный характер, что позволяет нам спрогнозировать поведение временного ряда. На рисунке 3 приведено сравнение одной из реализаций исходного сигнала $x(t)$ и спрогнозированного сигнала восстановленного $y(t)$ по трем гармоникам (6-й, 16-й и 21-й) с наибольшей амплитудой.

Однако следует учесть, что если сигнал имеет сложную форму, то с учетом наличия высокочастотных шумов на изображении трудно точно восстановить исходный сигнал, поэтому необходимо использовать фильтрацию спектра Фурье для увеличения точности прогнозирования. Так в приведенном примере величина средней абсолютной ошибки [5] для исходного $x(t)$ и прогнозируемого $x^*(t)$ временного ряда, рассчитанной по формуле

$$MAE = \frac{1}{N} \sum_{t=1}^N |x(t) - x^*(t)| \quad (6)$$

для приведенного примера составляют 0,25, или 25 %. Однако следует учесть, что значение восстановленного сигнала включает в себя и шумы.

Для устранения шумовой составляющей необходимо использовать фильтр низких частот (ФНЧ) с целью подавления высокочастотных компонент в спектре сигнала. В MATLAB реализована возможность конструирования фильтров различных типов и формы амплитудно-частотной характеристики (АЧХ). Пример АЧХ фильтра низких частот представлен ниже на рисунке 4.

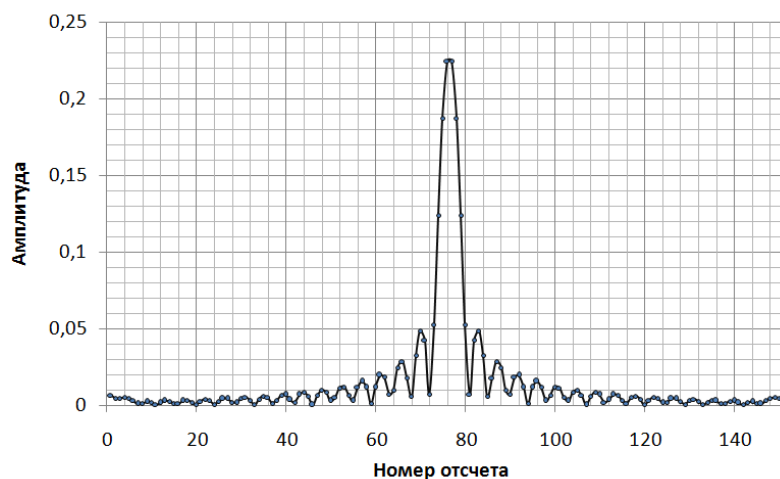


Рис. 4. АЧХ ФНЧ (окна преобразования)

Отфильтрованный сигнал также нетрудно восстановить. Амплитудный спектр фильтрованного сигнала является более гладким с ярко выраженными пиками на составляющих сигнал гармониках (рисунок 5).

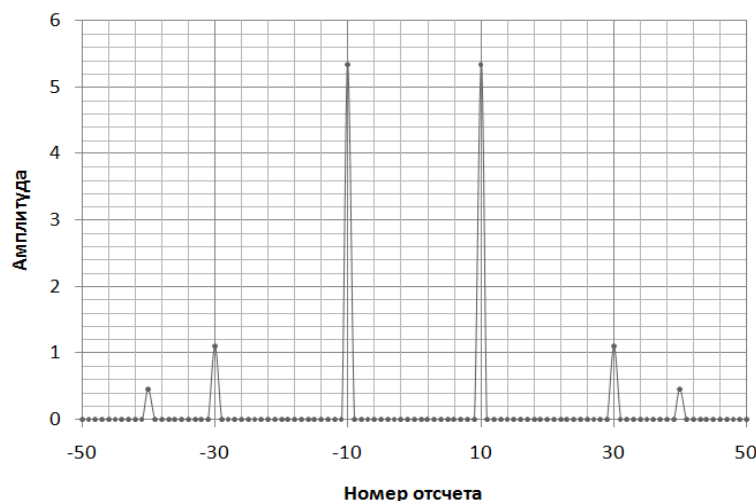


Рис. 5. Спектр отфильтрованного сигнала

При этом величина средней абсолютной ошибки между исходным временным рядом и рядом, восстановленным после фильтрации, составляет 0,28, или 28 %. Это означает, что шумовая составляющая отфильтрована и временной ряд еще больше отличается от исходного.

Кроме того, была рассчитана средняя абсолютная ошибка между прогнозируемым временным рядом без фильтрации и с фильтрацией, которая составляет 0,14, или 14 %. Именно этот показатель и влияет на точность прогноза после применения фильтрации, которая определяется выражением:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{z(t) - x^*(t)}{z(t)} \right| \cdot 100\%. \quad (7)$$

Для приведенного примера данный показатель составляет величину 93,8 %. В качестве $z(t)$ использовалась модель временного ряда без шумовой составляющей, что позволяет сделать вывод, что данная модель достаточно точно отражает действительные значения временного ряда.

В итоге с помощью Фурье-фильтрации можно получить более достоверный прогноз поведения стационарного временного ряда за счет использования фильтрации временного ряда в частотной области.

Список литературы

1. Грекова Т. И., Филатова Т. В. Построение трендовых моделей экономической системы // Вестник Томского государственного университета. – 2006. – № 292-2. – С. 294-297.
2. Буреева Н. Н., Ершова Ю. С. Влияние Фурье-фильтрации на построение прогноза экономических показателей с использованием аддитивной и мультипликативной моделей // Вестник Нижегородского университета им. Н. И. Лобачевского. – 2011. – № 3 (2). – С. 20-24.
3. Ляхманов Д. А., Соколова Э. С. Выделение детерминированной компоненты элементов временных рядов методом спектрального анализа // Вестник Нижегородского университета им. Н. И. Лобачевского. – 2011. – № 5 (1). – С. 181-184.
4. Цисарь И. Ф., Нейман В. Г. Компьютерное моделирование экономики. – М. : ДИАЛОГ-МИФИ, 2008. – 384 с.
5. Садовникова Н. А., Шмойлова Р. А. Анализ временных рядов и прогнозирование: учебное пособие. – М. : Московский государственный университет экономики, статистики и информатики, 2001. – 67 с.