КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ
COMPUTER SIMULATION HISTORY

УДК 004.942
DOI 10.52575/2687-0932-2024-51-2-402-412

О краткосрочном прогнозировании временных рядов на основе субполосных представлений

Жиляков Е.Г., Гайворонская Д.И., Прохоренко Е.И., Балабанова Т.Н.
Белгородский государственный национальный исследовательский университет
Россия, 308015, г. Белгород, ул. Победы, д. 85
E-mail: zhilyakov@bsu.edu.ru

Аннотация. Рассмотрена важная задача краткосрочного прогнозирования значений временных рядов на основе предыстории (экстраполяция) без априорного задания соотношений, определяющих зависимость будущих значений от предыдущих наблюдений. Показано, что в основу такой экстраполяции может быть положено понятие части энергии отрезка временного ряда, соответствующее некоторой субполосе области определения трансформанты Фурье (субполосное представление). Получены соотношения для вычисления прогноза и определены условия, при выполнении которых он будет безошибочным. Приведены оценки для среднеквадратических отклонений оценки будущего значения при наличии случайных ошибок измерений в виде белого шума.

Ключевые слова: временные ряды, краткосрочный прогноз, экстраполяция, субполосные представления

Для цитирования: Жиляков Е.Г., Гайворонская Д.И., Прохоренко Е.И., Балабанова Т.Н. 2024. О краткосрочном прогнозировании временных рядов на основе субполосных представлений. Экономика. Информатика, 51(2): 402–412. DOI 10.52575/2687-0932-2024-51-2-402-412

On Short-Term Forecasting of Time Series Based on Subband Representations

Evgeniy G. Zhilyakov, Diana I. Gaivoronskay, Ekaterina I. Prokhorenko, Tatiana N. Balabanova
Belgorod State National Research University
85 Pobedy St, Belgorod, 308015, Russia
E-mail: zhilyakov@bsu.edu.ru

Abstract. An important task of short-term forecasting of time series values based on prehistory (extrapolation) is considered without a priori setting the ratios that determine the dependence of future values on previous observations. It is shown that such an extrapolation can be based on the concept of the energy part of a time series segment corresponding to a certain subband of the domain of definition of the Fourier transform (subband representation). The relations for calculating the forecast are obtained and the conditions under which it will be error-free are determined. Estimates are given for the standard deviations of the future value estimation in the presence of random measurement errors in the form of white noise.

Keywords: time series, short-term forecast, extrapolation, subband representations

For citation: Zhilyakov E.G., Gaivoronskay D.I., Prokhorenko E.I., Balabanova T.N. 2024. On Short-Term Forecasting of Time Series Based on Subband Representations. Economics. Information technologies, 51(2): 402–412 (in Russian). DOI 10.52575/2687-0932-2024-51-2-402-412

Введение

Временным рядом $x_k = x(k\Delta t), k = 0,\pm1,\pm2,...$ будем называть результаты регистрации в эквидистантных с шагом Δt точках области определения заранее неизвестной функции времени $x(t)$, которая отражает свойства некоторого параметра наблюдаемого объекта, характеризующего его поведение. Следовательно, относительно будущих значений временных рядов имеется существенная априорная неопределенность. Вместе с тем без их предсказания (прогноза) часто оказывается невозможной адаптация к поведению наблюдаемых объектов реализуемых при этом процессов управления. Таким образом, предсказание будущих значений временных рядов является одной из основных процедур управления. Поэтому разработка методов прогнозирования относится к числу важных задач теории и практики обработки эмпирических данных.

Прогноз принято называть краткосрочным, если речь идет об оценке только нескольких (чаще всего одного) из ближайших к зарегистрированной предыстории будущих значений. Эта задача исследуется в работах многих авторов. Достаточно часто рассматривается прогноз на основе экстраполяции с использованием известной предыстории. При этом в основе прогнозирования используются так называемые инварианты, то есть некоторые характеристики временных рядов, мало изменяющиеся с течением времени.

Наиболее распространенной формой прогноза (экстраполяции) служит линейное предсказание (ЛП) в виде свертки

$$\widehat{x}_k = \sum_{i=1}^{p} \beta_i x_{k-i}, k > p, \tag{1}$$

где $\vec{\beta} = (\beta_1,...,\beta_p)'$ – вектор значений импульсной характеристики фильтра ЛП (штрих означает транспонирование); p – апертура фильтра.

В дальнейшем такие соотношения будем именовать моделями прогнозирования порядка p.

Очевидно, что прогноз (1) будет точным только тогда, когда значения временного ряда удовлетворяют разностному уравнению, коэффициенты которого известны. К таким временным рядам относятся аддитивные комбинации синусоид с различными частотами [Марпл, 1990]. Однако такие модели являются идеализацией, а значения их параметров (амплитуды, частоты и начальные фазы) заранее неизвестны и их надо оценивать, что снижает точность прогноза.

Адекватным подходом к прогнозированию последовательностей, удовлетворяющих однородным линейным разностным уравнениям, является использование метода главных компонент на основе сингулярного разложения [Уоткинс, 2009] матриц размерности $(p+1)*M$

$$F_k(p+1,M) = \{x_{n+m-1}\}, n = k-p-M-1,...,k-M; m = 1,...,M; p+1 \le M, \tag{2}$$

столбцы которых представляют собой последовательные отрезки временного ряда.

Этот метод прогнозирования в русскоязычной литературе часто именуется «гусеницей» [Галяндина, 2004], что связано со специфической ганкелевой формой матрицы с элементами (2). При его реализации определяются ортогональная матрица собственных векторов

$$Q_k = \{\vec{q}_1^{\,k},...,\vec{q}_{p+1}^{\,k}\} \tag{3}$$

матрицы (штрих означает транспонирование)

$$W_k = F_k(p+1,M)F_k^{'}(p+1,M), \tag{4}$$

так что имеет место

$$W_k Q_k = Q_k L_k, \tag{5}$$

где $L_k = diag(\lambda_1^k,...,\lambda_{p+1}^k)$ – диагональная матрица собственных чисел, которые предполагаются упорядоченными по убыванию

$$\lambda_1^k \ge ... \ge \lambda_{p+1}^k \ge 0. \tag{6}$$

Прогнозируемое значение определяется соотношением

$$q^k_{p+1,p+1}\hat{x}(k) = -\sum_{r=1}^{p} q^k_{p+1-r,p+1}x(k-r), \tag{7}$$

где вектор коэффициентов определяется компонентами собственного вектора матрицы (4), соответствующего минимальному собственному числу. Таким образом, в качестве инвариантов используются собственные векторы и числа матрицы (4), полученной на основе предыстории. Отметим, что этот подход до определенной степени аналогичен подходу Писаренко [Марпл, 1990] к спектральному оцениванию.

В общем случае из-за влияния различных факторов используемая предыстория и прогнозируемое значение не удовлетворяет разностному уравнению, что можно описать в виде аддитивного представления

$$\hat{x}(k) = x(k) + \varepsilon(k), \tag{8}$$

где погрешность прогноза $\varepsilon(k)$ определяется следующим соотношением:

$$\varepsilon(k) = -\sum_{r=0}^{p} q^k_{p+1-r,p+1}x(k-r)/q^k_{p+1,p+1}. \tag{9}$$

Из ортонормальности собственных векторов следует неравенство

$$|\, q^k_{p+1,p+1} | = (1-\sum_{r=1}^{p} q^{k2}_{p+1-r,p+1})^{1/2} < 1, \tag{10}$$

то есть знаменатель в (9) может быть очень мал. Следовательно, если сумма в правой части (9) отлична от нуля, то погрешность прогноза может быть велика.

Для описания неопределенностей чаще всего используются вероятностные модели. В рамках таких моделей в качестве адекватного критерия оптимальности (меры точности прогноза) используется среднеквадратическое отклонение ошибки прогноза. При этом правая часть (1) рассматривается как условное математическое ожидание процесса авторегрессии соответствующего порядка и параметров (AR(p)) [Бокс, Дженкинс, 1974]. Ясно, однако, что при обработке эмпирических данных параметры предполагаемых моделей их генерации необходимо оценивать и это увеличивает погрешность прогноза.

На практике для оценивания параметров линейного прогноза (1) широко используется принцип минимизации по набору параметров квадратичных функционалов вида

$$S_r(\vec{x}_N) = \sum_{k=p+1}^{N} (x_k - \sum_{i=1}^{r} \alpha_{kr}x_{k-i})^2, \tag{11}$$

где $\vec{x}_N = (x_1,...,x_N)'$ – обучающая выборка; r – величина апертуры фильтра ЛП (порядок модели AR(p))

$$1 \le r \le N/2,.... \tag{12}$$

Этот функционал при каждом значении апертуры минимизируется по значениям соответствующей импульсной характеристики фильтра ЛП и на оси r находится точка существенного уменьшения угла наклона к ней графика его минимальных значений, координата которой может быть принята за величину порядка. Полученная таким образом импульсная характеристика затем в качестве инварианта используется для прогнозирования на другой части последовательности эмпирических значений, если её вычисленные значения удовлетворяют условиям устойчивости модели [Андерсон, 1976; Кашьяп, 1983].

Ясно, что даже при справедливости гипотезы об адекватности модели AR(p) для достижения приемлемой точности оценивания её параметров требуется использовать обучающую выборку достаточно большой размерности. Вместе с тем в прогнозируемой выборке поведение значений локальных отрезков из p+1 отсчетов может плохо описываться линейной моделью, адаптированной к свойствам обучающей выборки. Это будет приводить к значительным ошибкам прогноза, по крайней мере отдельных значений.

Оптимальность в смысле минимума условного математического ожидания будущих значений при известной предыстории используется и в основе линейного прогнозирования векторных процессов по схеме Калмана [Сейдж, Мелс, 1976; Кузьмин, 1986]. Модель их генерации во многом является векторно-матричным обобщением уравнения авторегрессии. Очевидно, что и реализация фильтра Калмана предполагает наличие достаточно большого объема априорной информации о вероятностных свойствах обновляющей последовательности векторов и искажений при их регистрации.

Кажущаяся простота линейного прогноза способствовала его наибольшей распространенности, несмотря на затруднения с обоснованием адекватности формы (1), например, в случае изменчивости математического ожидания или других видов нестационарности. Кроме того, использование в качестве меры качества прогноза дисперсии ошибки позволяет гарантировать его оптимальность только в среднем на большом интервале наблюдений стационарных значений, тогда как целесообразно контролировать локальную ошибку особенно в нестационарном случае.

Отметим также, что по справедливому замечанию [Тьюки, 1981] априорное задание вида связи значений временного ряда равносильно навязыванию законов природе.

Перспективным представляется подход без оценивания параметров моделей генерации временных рядов с сохранением линейной формы их прогноза по небольшому количеству выборки ближайшей предыстории.

В данной работе в основе разработки такого подхода использованы субполосные представления [Жиляков, 2015; Жиляков, 2017], когда анализ данных осуществляется с позиций разбиения частотной полосы на субполосы, удовлетворяющие некоторым априорным требованиям.

Метод прогнозирования на основе субполосных представлений

Пусть $\vec{y}_k = (y_{k=p},...,y_k)'$ – вектор, состоящий из $p+1$ значений (компонент) некоторого временного ряда. Ему можно сопоставить трансформанту Фурье (спектр)

$$Y_k(v) = \sum_{i=1}^{p+1} y_{k-p-1+i} \exp[(-jv(i-1)), \tag{13}$$

которая, как легко видно, является периодической

$$Y_k(v + 2m\pi) = Y_k(v), \text{-}\,\pi \le v < \pi \,. \tag{14}$$

Справедливо равенство Парсеваля [Марпл, 1990]

$$\| \vec{y}_k \|^2 = \sum_{i=1}^{p+1} y^2_{k-p-1+i} = \int_{-\pi}^{\pi} |Y_k(v)|^2 \, dv/2\pi \,, \tag{15}$$

которое нетрудно представить в виде суммы

$$\| \vec{y}_k \|^2 = P_1(\vec{y}_k) + P_2(\vec{y}_k)\,, \tag{16}$$

где

$$P_1(\vec{y}_k) = \int_{-V}^{V} |Y_k(v)|^2 \, dv/2\pi, \; V < \pi\,; \tag{17}$$

$$P_2(\vec{y}_k) = \| \vec{y}_k \|^2 - P_1(\vec{y}_k)\,. \tag{18}$$

Таким образом, представляется естественным считать, что соотношение (17) определяет субполосную часть энергии (СЧЭ) рассматриваемого отрезка временного ряда, связанную с соответствующей субполосой частотной оси. Ниже будет показано, что использование этой субполосной характеристики отрезков временных рядов позволяет получить соотношение для линейного прогноза будущих значений.

В самом деле, подстановка определения (15) в представление (17) после несложных преобразований позволяет получить квадратичную форму для СЧЭ непосредственно во временной области

$$P_1(\vec{y}_k) = \vec{y}_k^{\,'} A_V \vec{y}_k\,, \tag{19}$$

где $A_V = \{a_{im}^V\}$. $i,m = 1,...,p+1$ – матрица с элементами

$$a_{im}^V = \int_{-v}^{V} \exp(-jv(i-m))dv/2\pi = \sin(V(i-m))/\pi(i-m)\,. \tag{20}$$

В дальнейшем матрица с элементами (20) называется субполосной. Свойства субполосных матриц достаточно подробно рассмотрены в работах [Жиляков, 2015; Жиляков, 2017].

Соотношение (19) нетрудно представить в следующем виде:

$$P_1(\vec{y}_k) = \sum_{i,m=1}^{p} a_{im}^V y_{k-p-1+i} y_{k-p-1+m} + \Delta P_1(\vec{y}_k)\,, \tag{21}$$

где имеется в виду часть энергии, связанная с последней компонентой рассматриваемого вектора

$$\Delta P_1(\vec{y}_k) = 2y_k \sum_{i=1}^{p} a_{p+1,i}^V y_{k-p-1+i} + a_{p+1,p+1}^V y_k^2\,. \tag{22}$$

Очевидно, что вклад последней компоненты в общую энергию вектора равен y_k^2. Только часть этого вклада будет влиять на СЧЭ рассматриваемой субполосы. Поэтому можно положить

$$\Delta P_1(\vec{y}_k) = c y_k^2\,, \tag{23}$$

где

$$|c| < 1\,. \tag{24}$$

В результате комбинирования (23), (22) и (20) нетрудно получить однородное квадратное уравнение, два корня которого определяются следующими соотношениями:

$$y_{1k} = 0\,; \tag{25}$$

$$y_{2k} = 2\sum_{i=1}^{p} a_{p+1,i}^V y_{k-p-1+i} /(c - V/\pi)\,. \tag{26}$$

Здесь имеется в виду, что знаменатель не равен нулю.

Легко также видеть, что подстановка корня (25) в правую часть определения (22) также дает ноль в правой части. Такое изменение СЧЭ (21) представляется маловероятным, если не все значения предыстории равны нулю. В этом случае нулю будет равна и правая часть соотношения (26), что представляется естественным значением прогноза в этих условиях. Поэтому в качестве основы уравнения порядка p для прогноза предлагается использовать правую часть (26), полагая

$$\hat{y}_k(c,V) = 2\sum_{i=1}^{p} a_{p+1,i}^V y_{k-p-1+i} /(c - V/\pi)\,. \tag{27}$$

Очевидно, что соотношение (27) по форме совпадает с (1), то есть оно определяет линейный прогноз.

Имея в виду определение (20), для точного прогноза должно выполняться равенство

$$y_k = \int_{-V}^{V} Y_{pk}(v) \exp(jv(m-1))dv/(\pi c - V)\,, \tag{28}$$

где

$$Y_{pk}(v) = \sum_{i=1}^{p} y_{k-p-1+i} \exp(-jv(i-1))\,. \tag{29}$$

Ясно, что при известном векторе $\vec{y}_k$ искомый параметр может быть вычислен из соотношения (28).

Таким образом, значение параметра c должно соответствовать ширине выбранной субполосы и частотным свойствам предыстории прогнозируемого значения временного ряда.

Очевидно, что точное прогнозирование на основе (27) достигается при выполнении следующего равенства для значений неизвестного параметра

$$c_k(V) - V/\pi = 2\sum_{i=1}^{p} a_{p+1,i}^{V} y_{k-p-1+i} / y_k \,. \tag{30}$$

Ввиду того, что в общем случае прогнозируемое значение неизвестно, вычислить отсюда точное значение параметра нельзя. Вместе с тем можно воспользоваться моделями локального поведения достаточно коротких отрезков временных рядов, чтобы получить уравнения для вычисления правой, а следовательно, и левой частей соотношения (30). Если эти модели описывают достаточно широкий класс локального поведения временных рядов на отрезке длительностью $p+1$, то вычисленную левую часть соотношения (30) можно использовать в качестве инварианта для всего ряда.

1. Пусть отрезок временного ряда имеет постоянные значения

$$\vec{y}_k = b(1,...,1)' \,. \tag{31}$$

Тогда из (30) получаем равенство, которое при неизменной субполосе определяет в качестве инварианта знаменатель в прогнозе (27), который позволяет сохранить значение постоянной составляющей предыстории

$$c(V) - V/\pi = 2\sum_{i=1}^{p} a_{p+1,i}^{V} \,. \tag{32}$$

2. Более широкий класс локального поведения временных рядов описывает линейная модель

$$y_{k-p-1+i} = b + (i-1)d, i = 1,..., p+1 \,. \tag{33}$$

где b и d – неизвестные параметры (числа). Легко получить модельное значение прогнозируемого отсчета

$$y_k = b + (p+1)d \,. \tag{34}$$

В соответствии с (30) для получения точного значения второго члена в правой части этого соотношения необходимо выполнить условие

$$(c(V) - V/\pi)(p+1) = 2\sum_{i=1}^{p} i a_{p+1,i} \,. \tag{35}$$

Таким образом, в случае адекватности модели (33) необходимо решать систему линейных алгебраических уравнений, определяемую соотношениями (32) и (35).

Если в (35) подставить правую часть (32), то с учетом определения (20) элементов субполосной матрицы нетрудно получить уравнение для ширины субполосы

$$\sum_{i=1}^{p} \sin(V(p+1-i)) = 0 \,, \tag{37}$$

которое с использованием формулы Эйлера для комплексной экспоненты преобразуется к виду

$$\sin(pV/2)\sin(V(p+1)/2) = 0 \,. \tag{38}$$

При этом необходимо учесть, что из периодичности трансформанты Фурье (14) и определения (17) следует неравенство для искомых корней уравнения (37)

$$0 < V_i < \pi, i = 1,.., p-1 \,. \tag{39}$$

Имея в виду равенства

$$\sin(i\pi) = 0, i = 0,.1,..,$$

получаем, что первый сомножитель в (38) равен нулю в точках

$$V_{1i} = 2i\pi / p, i < p/2 \,, \tag{40}$$

а для второго имеет место

$$V_{2i} = 2i\pi/(p+1), i < (p+1)/2. \quad (41)$$

Таким образом, совокупность соотношений (41) и (40) определяет корни уравнения (37), которые нужно вычислять с учетом неравенства (39).

При этом соотношение (32) дает

$$c(V_i) - V_i/\pi = 2\sum_{n=1}^{p} a_{p+1,n}^{V_i}, \quad (42)$$

а соотношение для прогноза (27) принимает окончательный вид

$$\hat{y}_k(c_i, V_i) = \sum_{n=1}^{p} a_{p+1,n}^{V_i} y_{k-p-1+n} / \sum_{m=1}^{p} a_{p+1,m}^{V_i}. \quad (43)$$

Для расширения класса полиномиальных моделей, адекватно с точки зрения точности прогноза описывающих локальное поведение коротких отрезков временных рядов, можно ряд $y_k, k = 1,...$получать на основе допустимых преобразований некоторого исходного, например, с помощью логарифмирования.

Привлекательным представляется использование разностей различного порядка

$$y_k = \Delta^{m-1}x_{k+1} - \Delta^{m-1}x_k, \quad (44)$$

где имеются в виду конечные разности соответствующего порядка некоторого исходного ряда x_k.

Важным частным случаем является ряд разностей первого порядка

$$y_k = x_{k+1} - x_k. \quad (45)$$

Очевидно, что модель прогноза значений исходного ряда при этом определяется соотношением

$$\hat{x}_{k+1} = x_k + \hat{y}_k. \quad (46)$$

Модель прогнозирования второго порядка на основе первых разностей

Речь идет о модели прогнозирования вида (46) с использованием соотношения (43), где

$$p = 2. \quad (47)$$

Целесообразность использования такой модели определяется, в частности, тем, что любые три значения исходного ряда можно точно аппроксимировать параболой, а любые две первых разностей удовлетворяют модели локального поведения (33), то есть аппроксимируются линейной моделью.

Единственным корнем уравнения (38) является (индекс опущен)

$$V = 2\pi/3. \quad (48)$$

Подстановка этого значения в (32) дает

$$2\sum_{n=1}^{2} a_{3,n} = 3^{1/2}/2\pi \approx 0{,}276, \quad (49)$$

а модель прогноза (43) преобразуется к виду

$$\hat{y}_k = 2y_{k-1} - y_{k-2}. \quad (50)$$

Подстановка правой части этого соотношения в (46) с учетом определения (45) дает модель прогноза без вычисления разностей

$$\hat{x}_{k+1} = 3x_k - 3x_{k-1} + x_{k-2}. \quad (51)$$

Представляет интерес оценка возможных ошибок прогнозирования на основе этой модели

$$r_{k+1} = |x_{k+1} - \hat{x}_{k+1}|. \quad (52)$$

При этом для коротких отрезков реальных временных рядов, получаемых с помощью дискретизации эмпирических функций с ограниченной энергией, которые имеют ограниченные производные любого порядка, адекватной математической основой представляется формула Тейлора. В самом деле для определенных в самом начале работы отсчетов некоторой непрерывной функции *x(t)* можно записать

$$x_{k-1} = x_{k-2} + x^{(1)}((k-2)\Delta t)\Delta t + x^{(2)}((k-2)\Delta t)(\Delta t)^2/2 + x^{(3)}(g_1)(\Delta t)^3/6\,; \quad (53)$$

$$x_k = x_{k-2} + 2x^{(1)}((k-2)\Delta t)\Delta t + 4x^{(2)}((k-2)\Delta t)(\Delta t)^2/2 + 8x^{(3)}(g_2)(\Delta t)^3/6\,; \quad (54)$$

$$x_{k+1} = x_{k-2} + 3x^{(1)}((k-2)\Delta t)\Delta t + 9x^{(2)}((k-2)\Delta t)(\Delta t)^2/2 + 27x^{(3)}(g_3)(\Delta t)^3/6\,, \quad (55)$$

где

$$g_i \in (x_{k-2}, x_{k-2} + i\Delta t), i = 1,2,3\,. \quad (56)$$

Подставив представления (53) и (54) в правую часть (51), нетрудно получить

$$\hat{x}_{k+1} = x_{k-2} + 3x^{(1)}((k-2)\Delta t)\Delta t + 9x^{(2)}((k-2)\Delta t)(\Delta t)^2/2 - (x^{(3)}(g_1) - 8x^{(3)}(g_2))(\Delta t)^3/2\,. \quad (57)$$

При сравнении с (55) легко видеть, что различия в правых частях определяются третьими производными. Поэтому определение (52) имеет представление

$$r_{k+1} = (\Delta t)^{|3}\,|\,9x^{(3)}(g_3) + x^{(3)}(g_1) - 8x^{(3)}(g_2)\,|\,/\,2\,. \quad (58)$$

Таким образом, при равенстве на рассматриваемом отрезке третьих производных исходной функции нулю прогнозирование будет осуществляться точно. Если третья производная везде равна некоторой константе *d*, то из соотношения (58) следует представление

$$r_{k+1} = (\Delta t)^{|3}\,|\,d\,|. \quad (59)$$

Очевидно, что погрешность существенно зависит от шага (частоты) дискретизации, и может служить основанием для его выбора.

Пусть теперь измерения отклоняются от некоторого тренда *x(t)*, так что отсчеты определяются соотношениями

$$\tilde{x}_{k-i} = x_{k-i} + \varepsilon_i\,, i = 0,1,2\,, \quad (60)$$

где ε_i – независимые центрированные случайные величины с дисперсией (*E* – символ оператора математического ожидания)

$$\sigma_\varepsilon^2 = E(\varepsilon_i^2) = const\,. \quad (61)$$

Тогда дисперсия вычисляемого на основе (60) прогноза (51), где неизвестные значения тренда можно считать постоянными, будет определяться соотношением

$$\sigma_{\hat{x}}^2 = 19\sigma_\varepsilon^2\,. \quad (62)$$

Иными словами, случайная погрешность прогноза буде пропорциональна величине

$$e = 4{,}35\sigma_\varepsilon\,. \quad (63)$$

Это соотношение показывает, что среднеквадратическое отклонение (СКО) случайной погрешности предлагаемого прогноза достаточно существенно превосходит СКО случайных отклонения исходных данных от предполагаемого тренда. Вместе с тем её значение устойчиво зависит от СКО случайных погрешностей регистрации исходных данных.

Еще один пример

Согласно соотношениям (40) и (41) при большем чем два порядке модели прогноза можно получить несколько возможных значений ширины используемой субполосы. Возникает вопрос о способе выбора из них в некотором смысле наиболее приемлемого.

Пусть порядок модели выбран равным

$$p = 4\,, \quad (64)$$

что является достаточно большим значением.

Тогда согласно (40) и (41) совокупность корней уравнения (38) (учетом (39)) имеет вид

$$V_1 = \pi/3\,;\ V_2 = 2\pi/5\,;\ V_3 = 2\pi/3\,..V_4 = 4\pi/5\,. \quad (65)$$

Отметим, что количество корней совпадает с порядком модели прогнозирования минус единица.

Рассмотрим подробнее прогнозирование при выборе первого из корней. В соответствии с (32) получаем

$$c_1 - V_1 / \pi = 1{,}45(3)^{1/2} / 2\pi\,; \tag{66}$$

$$\hat{x}_{k+1} = (2{,}45x_k - 0{,}5x_{k-1} - 0{,}5x_{k-2} - 0{,}25x_{k-3} + 0{,}45x_{k-4} - 0{,}2x_{k-5})/1{,}45\,. \tag{67}$$

В условиях представления (60) и свойства (61) нетрудно получить соотношение для среднеквадратического отклонения прогноза, обусловленного наличием случайной ошибки измерения

$$e = 1{,}73\sigma_\varepsilon\,. \tag{68}$$

Нетрудно видеть, что правая часть здесь гораздо меньше, чем у (63). Этот результат соответствует соотношению (28), которое показывает, что уменьшение (в данном случае в два раза) ширины субполосы должно приводить к уменьшению влияния случайной погрешности типа белого шума.

Если выбрать третий из корней совокупности (65), что совпадает с рассмотренным ранее случаем второго порядка

$$V = 2\pi/3\,, \tag{69}$$

то получим

$$c_3 - V_3 / \pi = 0{,}55(3)^{1/2} / 2\pi\,; \tag{70}$$

$$\hat{x}_{k+1} = (1{,}55x_k - 1{,}5x_{k-1} + 0{,}5x_{k-2} + 0{,}25x_{k-3} - 0{,}45x_{k-4} + 0{,}2x_{k-5})/0{,}55\,; \tag{71}$$

$$e = 3{,}90\sigma_\varepsilon\,. \tag{72}$$

Легко видеть, что среднеквадратическое отклонение выросло пропорционально ширине субполосы, однако остается меньше, чем в (63).

Этот пример показывает, что при выборе порядка модели прогноза целесообразно провести анализ влияния выбора ширины субполосы. При выборе порядка модели прогнозирования важное значение имеет точность выполнения условия (33) на интервале прогнозирования. Очевидно, что чем меньше порядок, тем обоснованнее считать, что оно выполняется.

Заключение

Прогнозирование временных рядов является одной из важнейших проблем управления различными процессами и объектами на основе адаптации к конкретным условиям. При этом особое значение имеет краткосрочное прогнозирование на основе небольшого количества значений ближайшей предыстории, так как это позволяет реагировать на возможные изменения свойств временных рядов. Представляется важным не использовать априорного постулирования формы зависимости будущего от прошлого. В основе большинства известных методов используется именно такой подход. Одним из исключений служит так называемый метод «гусеница», в котором неявно используется оценивание матрицы ковариаций и проектирование отрезка временного ряда на собственный вектор этой матрицы, который соответствует минимальному собственному числу. Одним из недостатков метода является риск получения очень большой погрешности прогноза при наличии шумов измерений. Кроме того, метод применим, когда минимальное собственное число очень мало.

В рамках данной работы предложен метод прогнозирования на основе субполосных представлений, который устойчив к воздействиям шумов измерений.

Список литературы

Марпл С.Л.(мл.). 1990. Цифровой спектральный анализ и его приложения. М.: Мир,

Уоткинс Д.С. 2009. Основы матричных вычислений: пер. со второго англ. изд. Уоткинс Д.С.; пер. с англ. Кондрашов В.Е., Королев С.Б. М.: БИНОМ. Лаборатория знаний, 664 с.

Галяндина Н.Э. 2004. Метод «гусеница» – SSA: прогноз временных рядов. С.-Петербургский государственный университет.

Бокс Дж., Дженкинс Г. 1976. Анализ временных рядов: прогноз и управление М.: Мир. Вып. 1. 1974. 406 с.

Андерсон, Теодор Вибурн. Статистический анализ временных рядов: пер. с англ. И.Г. Журбенко и В.П. Носко; под ред. Ю.К. Беляева. Москва: Мир, 755 с.

Кашьяп Р.Л. 1983. Построение динамических стохастических моделей по экспериментальным данным: пер. с англ. Т. И. Дубенко и др. Москва: Наука, 383 с.

Сейдж Э., Мелс Дж. 1976. Теория оценивания и ее применение в связи и управлении: пер. с англ. под ред. проф. Б.Р. Левина. 495 с.

Кузьмин С.З. 1986. Основы проектирования систем цифровой обработки радиолокационной информации. М.: Радио и связь, 352 с.

Коновалов А.А. 2014. Основы траекторной обработки радиолокационной информации. С.-Петербург, изд-во СПб ЭТУ «ЛЭТИ», 179 с.

Тьюки Д.У. 1981. Анализ результатов наблюдений: разведочный анализ: пер. с англ. А.Ф. Кушнира [и др.] Москва: Мир, 693 с.

Жиляков Е.Г. 2015. Оптимальные субполосные методы анализа и синтеза сигналов конечной длительности. Автомат. и телемех., №4: 52–66.

Жиляков Е.Г. 2017. Построение трендов отрезков временных рядов. Автомат. и телемех., № 3: 80–95.

References

Marpl S.L. (ml.) 1990. Cifrovoj spektral'nyj analiz i ego prilozheniya [Digital spectral analysis and its applications]. M.: Mir.

David S. 2009. Watkins Fundamentals of Matrix Computations; per. s angl. Kondrashov V. E., Korolev S. B, M.: BINOM. Laboratoriya znanij, 664 p.

Galyandina N.E. 2004. Metod «gusenica» – SSA: prognoz vremennyh ryadov [The caterpillar - SSA method: time series prediction]. S.-Peterburgskij gosudarsitvennyj universitet.

George E.P. 1974. Box, Gwilym M. Jenkins Time Series Analysis: Forecasting and Control - M.: Mir. Vyp. 1 406 p.

Theodore W. Anderson. 1976. The Statistical Analysis of Time Series: per. s angl. I.G. ZHurbenko i V.P. Nosko; pod red. YU.K. Belyaeva Moskva: Mir, 755 p.

R.L. Kashyap, A. Ramachandra Rao Dynamic stochastic models from empirical data: per. s angl. T.I. Dubenko i dr. Moskva: Nauka, 1983, 383 p.

Sage A.P., Melse J.L. 1972. Estimation theory with application to communication and control. N.Y.: McGraw-Hill Book Co., 540 p. (Russ. ed.: Sage, A., Melse, J. Teoriya otsenivaniya i ee primenenie v svyazi i upravlenii. Moscow: Svyaz', 1976. 495 p.).

Kuzmin S.Z. 1986. Osnovy proektirovaniya sistem cifrovoj obrabotki radiolokacionnoj informacii [Fundamentals of designing digital radar information processing systems] M.: Radio i svyaz', 352 p.

Konovalov A.A. 2014. Osnovy traektornoj obrabotki radiolokacionnoj informacii [Fundamentals of trajectory processing of radar information]. S.-Peterburg, izd-vo SPb ETU»LETI» - 179 p.

John W. Tukey. 1981. Exploratory data analysis: per. s angl. A. F. Kushnira [i dr.] Moskva: Mir, 693 p.

Zhilyakov E.G. 2015. Optimal'nye subpolosnye metody analiza i sinteza signalov konechnoj dlitel'nosti [Optimal subband methods for analysis and synthesis of finite duration signals]. Avtomat. i telemekh., №4: 52–66

Zhilyakov E.G. 2017. Postroenie trendov otrezkov vremennyh ryadov [Building trends of time series segments]. Avtomat. i telemekh., № 3, 80–95.

Конфликт интересов: о потенциальном конфликте интересов не сообщалось.
Conflict of interest: no potential conflict of interest related to this article was reported.

Поступила в редакцию 26.04.2024 Received April 26, 2024
Поступила после рецензирования 03.06.2024 Revised June 03, 2024
Принята к публикации 05.06.2024 Accepted June 05, 2024

ИНФОРМАЦИЯ ОБ АВТОРАХ

Жиляков Евгений Георгиевич, доктор технических наук, профессор, профессор кафедры информационно-телекоммуникационных систем и технологий института инженерных и цифровых технологий, Белгородский государственный национальный исследовательский университет, Белгород, Россия

Гайворонская Диана Игоревна, кандидат технических наук, доцент кафедры информационно-телекоммуникационных систем и технологий института инженерных и цифровых технологий, Белгородский государственный национальный исследовательский университет, Белгород, Россия

Прохоренко Екатерина Ивановна, кандидат технических наук, доцент, доцент кафедры информационно-телекоммуникационных систем и технологий института инженерных и цифровых технологий, Белгородский государственный национальный исследовательский университет, Белгород, Россия

Балабанова Татьяна Николаевна, кандидат технических наук, доцент, доцент кафедры информационно-телекоммуникационных систем и технологий института инженерных и цифровых технологий, Белгородский государственный национальный исследовательский университет, Белгород, Россия

INFORMATION ABOUT THE AUTHORS

Evgeniy G. Zhilyakov, Doctor of Technical Sciences, Professor, Professor of the Department of Information and Telecommunication Systems and Technologies, Belgorod State National Research University, Belgorod, Russia

Diana I. Gaivoronskay, Candidare of Technical Sciences, Associate Professor of the Department of Information and Telecommunication Systems and Technologies, Belgorod State National Research University, Belgorod, Russia

Ekaterina I. Prokhorenko, Candidare of Technical Sciences, Associate Professor, Associate Professor of the Department of Information and Telecommunication Systems and Technologies, Belgorod State National Research University, Belgorod, Russia

Tatiana N. Balabanova, Candidare of Technical Sciences, Associate Professor, Associate Professor of the Department of Information and Telecommunication Systems and Technologies, Belgorod State National Research University, Belgorod, Russia