Journal of Siberian Federal University. Engineering & Technologies 3 (2014 7) 359-363
~ ~ ~

УДК 519.6

Анализ применения комбинированных моделей при краткосрочном прогнозировании временных рядов

А.А. Пьяных*
Сибирский федеральный университет,
Россия, 660041, Красноярск, пр. Свободный, 79

Received 30.12.2013, received in revised form 15.03.2014, accepted 25.04.2014

Доказано на примере модели Брауна нулевого порядка, что точность прогнозирования повышается при подборе длины адаптационного окна, а также при увеличении пределов допустимых значений параметра адаптации. На основе результатов вычислительного эксперимента показано, что точность прогнозирования увеличивается с применением алгоритмов комбинирования моделей.

Ключевые слова: временной ряд, прогнозирование, адаптивная модель.

Введение

Временные ряды, как правило, возникают в результате измерения некоторого показателя. Это могут быть показатели как технических систем, так и природных, социальных, экономических и др.

Прогнозирование и анализ временных рядов находят широкое применение в сфере энергетики. Так, например, могут рассматриваться следующие характеристики: цена реализуемой электроэнергии, объем производства электроэнергии, объем потребления электроэнергии, объемы использования сырья, эффективность работы электростанций, гидроэлектростанций и атомных электростанций, энергоэффективность использования топлива, тепловой энергии (горячая вода, водяной пар, хладагенты), воды, суммарная мощность источников теплоснабжения, число источников теплоснабжения, шумовое воздействие оборудования и мн. др.

Временные ряды в энергетике можно использовать для интерпретации полученных данных. С их помощью можно определять тенденцию изменения показателя, например, растут ли объемы производства энергии или, наоборот, снижаются. Временные ряды наглядно показывают колебания рассматриваемой характеристики, например, на графике рядов можно увидеть суточные и сезонные колебания потребления электроэнергии.

* Corresponding author E-mail address: t.pyanykh@gmail.com

Постановка задачи исследования

Задача прогнозирования временного ряда заключается в следующем. Пусть задан временной ряд: $X = \{x_t : t \in T, x_t \in R\}$, $T = \{1,2,\ldots N\}$. Необходимо найти предсказанные значения величин $x_{t+\tau|t}$ временного ряда, $\tau>0$, по предыдущим наблюдениям $\{x_1,x_2,\ldots x_t\}$, что сводится к нахождению детерминированной функции t аргументов, значения которой можно было бы принимать в качестве $x_{t+\tau|t}$, т. е.

$$\hat{x}_{t+\tau|t} = f_{t,\tau}(x_1, x_2, ..., x_t), t \in \{t_1, t_2, ..., t_N\}, \tau \in \{1,2,..., D\}, \tag{1}$$

где D – горизонт прогнозирования. В данной работе $D = 1$, что соответствует краткосрочным прогнозам. Среди всевозможных функций (1) представляет интерес та, которая дает наилучшее предсказание. В анализе временных рядов принято считать наилучшим такое предсказание (прогнозирование), которое обеспечивает минимальную среднеквадратическую ошибку. Иными словами, функция (1) должна быть решением следующего экстремального уравнения:

$$I = \frac{1}{D}\sum_{i=1}^{D}\left(x_{t+i} - \hat{x}_{t+i|t}\right)^2 \to \min_{f(\cdot)} \cdot \tag{2}$$

Алгоритмы комбинирования моделей

1. Адаптивная селективная композиция моделей. Суть этой модели заключается в следующем. На каждом шаге по нескольким базовым моделям определяют прогнозные значения, затем сравнивают их с фактическими, после чего модель, которая показала лучшие результаты, используется для нахождения новых прогнозных значений. На следующем шаге процедура повторяется [1]. Таким образом, прогноз на τ шагов определяется следующим образом:

$$\hat{x}_{t+\tau|t} := \hat{x}_{t+\tau|t}^{\left(j_t^*\right)}, \tau > 0, \tag{3}$$

где $\hat{x}_{t+\tau|t}^{\left(j_t^*\right)}$ – прогноз модели под номером j_t^* в момент времени t на τ шагов. Номер модели в момент времени t определяется следующим образом:

$$j_t^* = \arg \min_{j=1,...,k} \tilde{\delta}_{t,j}, \tag{4}$$

здесь K – количество моделей базового набора; $\tilde{\delta}_{t,j}$ – экспоненциально сглаженная средняя квадратическая ошибка модели под номером j в момент времени t:

$$\tilde{\delta}_{t,j} = \gamma\delta_{t,j} + (1-\gamma)\tilde{\delta}_{t-1,j}, \tag{5}$$

$$\delta_{t,j} = \frac{1}{K}\sum_{i=1}^{K}\left(x_{t,j} - \hat{x}_{t,j}\right)^2. \tag{6}$$

Необходимо отметить, что использование адаптивной селективной модели эффективно, когда базовые модели существенно различаются [1].

2. Адаптивная гибридная композиция моделей. Для тех случаев, когда в адаптивную комбинированную модель входят модели, дающие сравнительно близкие результаты, и селекция затруднена, можно использовать гибридную адаптивную композицию моделей, прогноз по которой является взвешенной суммой прогнозов, полученных по входящим в нее предикторам. Веса прогнозов $\omega_{t,j}$ предлагается брать адаптивными:

$$\omega_{t,j} = \frac{\left(\tilde{\delta}_{t,j}\right)^{-1}}{\sum_{j=1}^{K}\left(\tilde{\delta}_{t,j}\right)^{-1}}. \tag{7}$$

Предиктор адаптивной гибридной модели определяется следующим образом:

$$\hat{x}_{t+\tau|t} = \sum_{j=1}^{K} \omega_{t,j} \hat{x}^{(j)}_{t+\tau|t}, \tau > 0. \tag{8}$$

В отличие от модели, рассмотренной ранее, гибридная модель осуществляет переключение с одной модели на другую более плавно со множеством промежуточных положений. Результаты, полученные различными моделями, могут случайно приближаться к реальному процессу и кратковременно давать хорошие прогнозы, что может привести к увеличению их весовых коэффициентов в адаптивной гибридной модели. Вследствие этого могут появиться ошибки, которые могут снизить эффективность комбинирования моделей [1]. Для уменьшения такого эффекта была использована экспоненциально сглаженная ошибка прогнозов $\tilde{\delta}$ при вычислении весовых коэффициентов по формуле (7).

Совершенствование модели Брауна нулевого порядка

В классической модели Брауна нулевого порядка предполагается, что область допустимых значений параметра адаптации лежит в пределах от 0 до 1. В работе [2] показано, что увеличение предельных значений параметра адаптации повышает точность прогнозов.

Одним из интересных вопросов является исследование зависимости между длиной адаптационного окна (количество членов ряда, используемых для прогноза на текущем временном шаге) и точностью прогнозов. Помимо коэффициента экспоненциального сглаживания также целесообразно подбирать на каждом шаге значение длины адаптационного окна. Данный параметр может существенно зависеть от исследуемого временного ряда: для рядов с длинной актуальной частью он будет принимать большие значения, а для рядов с короткой актуальной частью – меньшие.

Численный пример

Для прогнозирования использовали модели Брауна нулевого, первого и второго порядков, модель экспоненциального роста, а также модель Хольта [1, 3]. Тестирование алгоритмов краткосрочных прогнозов было проведено на курсе доллара США [4] (с 11.01.2012 по 25.05.2013) с применением различных значений длины адаптационного окна. Точность моделей прогнозирования оценивали по значениям среднеквадратической ошибки прогнозов.

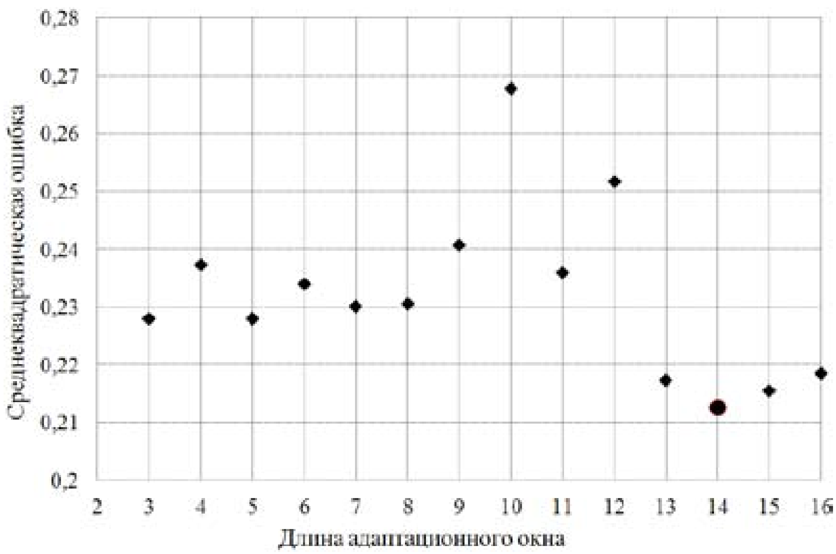

Рис. 1. Зависимость среднеквадратической ошибки прогнозов, полученных методом Брауна нулевого порядка, от длины адаптационного окна

На рис. 1 представлена зависимость значения среднеквадратической ошибки от длины адаптационного окна в модели Брауна нулевого порядка, в которой допустимые значения параметра адаптации принимались в пределах от 0 до 3. Как видим, при длине адаптационного окна, равной 14, модель Брауна нулевого порядка показывала самые точные прогнозы.

В табл. 1 сведены результаты прогнозов, получаемых различными моделями. Выявлено, что применение селективной композиции моделей позволяет получать наименьшую среднеквадратическую ошибку прогнозов (0,16 руб.). В целом композиция моделей увеличивает точность прогнозов.

Программная реализация алгоритмов

Программное обеспечение было разработано на языке Java с использованием среды визуального программирования Eclipse Indigo и функционирует в операционной системе из семейства Windows XP/Vista/7/8. Графический интерфейс программы был разработан с использованием библиотеки Swing, разработанной компанией Sun Microsystems.

При проектировании программных алгоритмов прогнозирования временных рядов был применен объектно-ориентированный подход.

Заключение

В работе выполнено совершенствование модели Брауна нулевого порядка за счет подбора длины адаптационного окна, а также увеличения пределов допустимых значений параметра адаптации, что позволяет получать значительно более точные результаты прогнозирования.

Показано, что прогнозные расчеты с помощью комбинирования моделей дают более точные результаты по сравнению с отдельными моделями.

Таблица 1. Результаты прогнозирования

Наименование модели	Среднеквадратическая ошибка, руб.	Максимальная ошибка, руб.	Минимальная ошибка, руб.
Модель Брауна нулевого порядка	0,229	1,174	6,4E-07
Модель Брауна первого порядка	0,2565	1,151	6,4E-07
Модель Брауна второго порядка	0,281	1,458	1,6E-07
Модель экспоненциального роста	0,302	1,689	4,9E-07
Модель Хольта	0,302	1,618	4,9E-07
Селективная композиция моделей	0,16	0,867	1,6E-07
Гибридная композиция моделей	0,202	1,133	1,12E-07
Минимальные значения	0,16	0,867	1,12E-07

Список литературы

[1] *Лукашин Ю.П.* Адаптивные методы краткосрочного прогнозирования временных рядов: учеб. пособие. М.: Финансы и статистика, 2003. 416 с.

[2] *Светуньков С.Г.* // Известия Санкт-Петербургского государственного университета экономики и финансов. 2002. № 3. С. 94-107.

[3] *Wang S.* PhD dissertation. Georgia Institute of Technology, Georgia, 2006.

[4] [Электронный ресурс]. – Режим доступа: http://www.cbr.ru

Analysis of Application of the Combined Models at Short-Term Forecasting of Time Series

Artem A. Pyanykh
Siberian Federal University
79 Svobodny, Krasnoyarsk, 660041, Russia

It is proved on an example of model of Brown of a zero-order, that accuracy of forecasting increases at selection of length of an adaptable window, and also increase limits of admissible values of adaptation parameter. On the basis of results of computing experiment it is shown, that accuracy of forecasting raises with application of algorithms of a combination of models.

Keywords: time series, forecasting, adaptive model.