Локальные методы прогнозирования с выбором преобразования*

С. В. Цыганова
schiavoni@mail.com
Московский физико-технический институт, ФУПМ, каф. «Интеллектуальные системы»

В работе описан алгоритм локального прогнозирования с учетом преобразований, позволяющий выявить похожие во введеной метрике интервалы временного ряда. Рассмотренно понятие инвариантных преобразований, их обнаружение и выбор наиболее подходящих для решения задачи прогнозирования. Работа алгоритма проиллюстрирована на данных потребления электроэнергии и на синтетических данных.

Ключевые слова: *локальное прогнозирование, функция расстояния, временной ряд, инвариантное преобразование*

Methods of local forecasting with transormation accounting*

S. V. Tsyganova
Moscow Institute of Physics and Technology

This paper considers the algorithm of local forecasting with transformation, which reveals similar intervals of time series in introduced metrics. A conception of invariant transformations is considered, and also the choice of the most suitable for forecasting problem. The work is illustrated by the data of energy consumption and synthetic data.

Keywords: *local forecasting, distance function, time series, invariant conversion.*

Задачи прогнозирования временных рядов имеют множество приложений в различных областях, таких как экономика, физика, медицина. Их решением является прогноз на недалекое будущее по уже известным значениям прогнозируемого ряда в предыдущие моменты времени. Данная работа посвящена методу локального прогнозирования временных рядов. Для построения прогноза используются только те части временного ряда, которые близки к конечному отрезку всего временного ряда. Близкими считаются те отрезки временных рядов, функция близости для которых мала. Для определения близких отрезков в работе исследуется линейное преобразование (сжатие, сдвиг), инварианты преобразований и функция «близости» отрезков временных рядов, которая будет являться одним из критериев адекватной работы построенного алгоритма прогнозирования. Общий локальный метод прогнозирования основан на идеях, описанных в работе Дж. Макнеймса [1] и Ю.И. Журавлева [2].

Для нахождения близких интервалов использован метод «ближайшего соседа», успешно применяемый к широкому классу прикладных задач, таких как прогнозирование объемов продаж, прогнозирование цен на электроэнергию, постановка диагноза по биоритмам человека.

Сложности при построении алгоритма – это учет пропусков в предоставленных данных. В данной работе считается, что данные представлены без пробелов.

Научный руководитель В. В. Стрижов

Проверка алгоритма будет производиться при помощи скользящего контроля, т.е. прогноз будет сравниваться с реальными значениями.

Вся работа разделена на четыре главы. Первая глава – это математическая постановка задачи. Во второй главе описывается алгоритм преобразования и прогнозирования с некоторыми математическими выкладками. Третья глава – вычислительный эксперимент для двух временных рядов (синтетического и потребления энергии [6]) и исследование эффективности алгоритма. В последней главе сформулирован общий вывод.

Постановка задачи

Будем рассматривать одномерные временные ряды — ряды, в которых каждому моменту времени сопоставляется вещественное число.

$$\{t_1, t_2, \dots, t_n\} \to \{x_1, x_2, \dots, x_n\}$$

Требуется предсказать следующие l значений последовательности $\{x_{n+1}, x_{n+2}, \dots, x_{n+l}\}$, которые будут определяться значением предыстроии $\{x_{n-L+1}, x_{n-L+2}, \dots, x_n\}$ длины L. Для этого необходимо выполнить следующий алгоритм:

1. Выделить во всем временном ряде вектора длины

$$r : r_{min}, \dots, r_{max},$$

которые после линейных преобразований A (сжатие,сдвиг) похожи на предысторию $S = \{x_{n-L+1}, x_{n-L+2}, \dots, x_n\}$.

2. Найти и исследовать инварианты преобразований между двумя близкими векторами временного ряда и с их помощью найти те самые «похожие» вектора.

3. Критерий близости играет функция близости двух векторов **a**, **b** — в данной работе это взвешенная метрика Евклида:

$$D_{WE}(a, b) = \sqrt{(a-b)^T \Lambda^2 (a-b)}.$$

4. Задача формулируется следующим образом:

$$\mathrm{dist}(A(x_{k-r+1}, x_{k-r+2}, \dots, x_k), (x_{n-L+1}, x_{n-L+2}, \dots, x_n)) \to min,$$

где $A(x_{k-r+1}, x_{k-r+2}, \dots, x_k)$ – преобразованный близкий вектор, а $(x_{n-L+1}, x_{n-L+2}, \dots, x_n)$ – вектор предыстории.

5. Для отыскания k близких векторов используем метод k ближайших соседей. Пусть

$$\{A_1(x_{i_1-r+1}, \dots, x_{i_1}), \dots, A_k(x_{i_k-r+1}, \dots, x_{i_k})\} - -$$

это k ближайших соседей для предыстории $\{x_{n-L+1}, x_{n-L+2}, \dots, x_n\}$.

Прогноз вычисляется как среднее k векторов:

$$\{A_1(x_{i_1+1}, \dots, x_{i_1+l}), \dots, A_k(x_{i_k+1}, \dots, x_{i_k+l})\},$$

где среднее вычисляется как взвешенное среднее арифметическое:

$$(x_{i_1+1}, \dots, x_{i_1+l}) = \frac{\sum_{j=1}^{k} w_j A_j(x_{i_j+1}, \dots, x_{i_j+l})}{\sum_{j=1}^{k} w_j},$$

$$w_j = \left(1 - \frac{d_{ij}^2}{d_{i_{k+1}}^2}\right)^2,$$

где $d_{i_{k+1}}^2$ – расстояние до $k+1$ ближайшего соседа.

Описание алгоритма

Алгоритм включает в себя следующие этапы:

1. Нахождение преобразования φ по оси OX и выбор потенциальных соседей – векторов $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$. Для этого во всем временном ряде X находятся точки экстремумов. Далее находятся такие векторы, чтобы точки экстремальных значений этих векторов до точек экстремальных значений предыстории имели минимальное расстояние в выбранной метрике. Максимальное отклонение, допускаемое алгоритмом – заданный параметр ε. Из этого условия для каждого i-го потенциального соседа находится коэффициент a_{0_i} растяжения по OX, а выбранные векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ становятся потенциальными соседями. Последующие значения – $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_m$ – потенциальным прогнозом в зависимости от близости соседа. Количество потенциальных соседей прямопропорционально параметру ε: чем меньше параметр, тем меньше потенциальных соседей выделяет алгоритм.

Таким образом мы находим соседние векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ для предыстории $\{x_{n-L+1}, x_{n-L+2}, \ldots, x_n\}$ у временных рядов не только с постоянным, но и с изменяющимся периодом.

2. Для каждого потенциального соседа минимизируется функция близости и находятся коэффициенты b_{0_i} растяжения по OY для каждого близкого вектора. Пусть $\mathbf{y}$ – данный вектор временного ряда, а $\mathbf{x}$ – вектор временного ряда, который необходимо преобразовать по оси OY, чтобы получить наиболее близкий к $\mathbf{y}$ вектор. Тогда требуется найти минимум следующей функции:

$$F = \sum_{t=1}^{l} (y_t - (a + bx_t))^2.$$

Необходимые условия экстремума (более подробно смотри [3]):

$$\begin{cases} \frac{dF}{da} = -2\sum_{t=1}^{l}(y_t - a - bx_t) = 0, \\ \frac{dF}{db} = -2\sum_{t=1}^{l} x_t(y_t - a - bx_t) = 0. \end{cases}$$

Раскроем скобки и получим систему уравнений:

$$\begin{cases} al + b\sum_{t=1}^{l} x_t = \sum_{t=1}^{l} y_t, \\ a\sum_{t=1}^{l} x_t + b\sum_{t=1}^{l} x_t^2 = \sum_{t=1}^{l} x_t y_t. \end{cases}$$

Решениями системы являются:

$$\begin{cases} b = \frac{l\sum_{t=1}^{l} x_t y_t - (\sum_{t=1}^{l} x_t)(\sum_{t=1}^{l} y_t)}{n\sum_{t=1}^{l} x_t^2 - (\sum_{t=1}^{l} x_t)^2}, \\ a = \frac{1}{l}\sum_{t=1}^{l} y_t - \frac{1}{l}\sum_{t=1}^{l} x_t b. \end{cases}$$

3. Соседи сортируются по значению функции близости. Далее выбираются k самых близких (k – заданное число). Их потенциальный прогноз усредняется (чем меньше значение функции близости для соседа, тем больший вклад дает k-й потенциальный прогноз в усреднение).

4. Построение прогноза и вычисление ошибки с помощью скользящего контроля. Для сравнения всех прогнозов в работе используется суммарная абсолютная ошибка отклонения прогноза от действительных значений. Обозначим $\mathbf{f} = (f_{n+1}, \ldots, f_{n+l})$ – точное значение временного ряда и $\tilde{\mathbf{f}} = (\tilde{f}_{n+1}, \ldots, \tilde{f}_{n+l})$ – полученный алгоритмом прогноз. Тогда качество алгоритмов сравнивается при помощи следующей величины:

$$E = \frac{1}{l} \sum_{j=1}^{l} \left| f_{n+j} - \tilde{f}_{n+j} \right|.$$

Этот функционал ошибки зависит только от абсолютного отклонения прогноза от точных значений временного ряда и не зависит от их величины.

Вычислительный эксперимент и эффективность

Данный метод прогнозирования предназначен для прогнозирования временных рядов с переменным периодом, что дает алгоритму преимущества перед алгоритмом, использующим преобразования сжатия и сдвига по оси OY и описанным в работе В. Федоровой [4].

Рассмотрим следующий пример – спрогнозируем модельный временной ряд :

$$y = sin(x^2)$$

Сравним работу следующих двух алгоритмов – с использованием преобразования по OX и без:

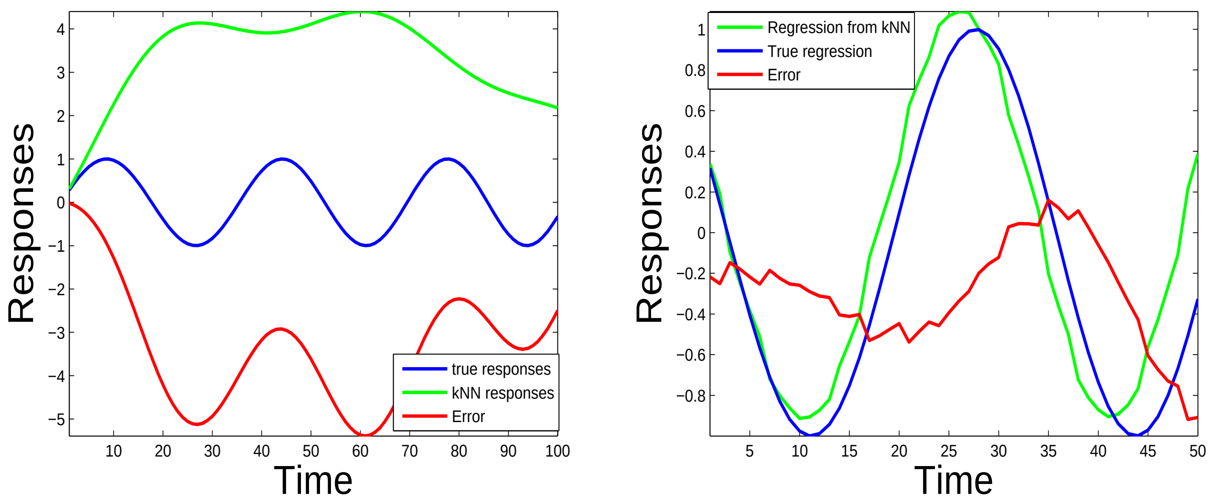

Рис. 1. Прогноз с использованием преобразования по OX (справа) и без (слева)

Теперь сравним работу алгоритмов на реальных данных – будем прогнозировать потребление энергии на один день вперед. График потребления электроэнергии за последние несколько дней (625 временных точек) выглядит следующим образом:

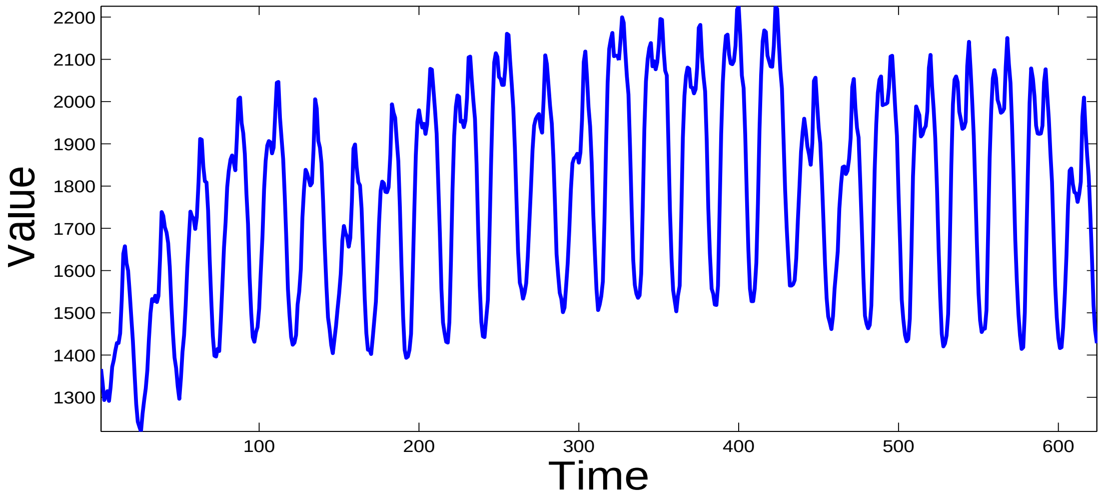

Рис. 2. график потребления электроэнергии за несколько дней

Результат работы алгоритмов:

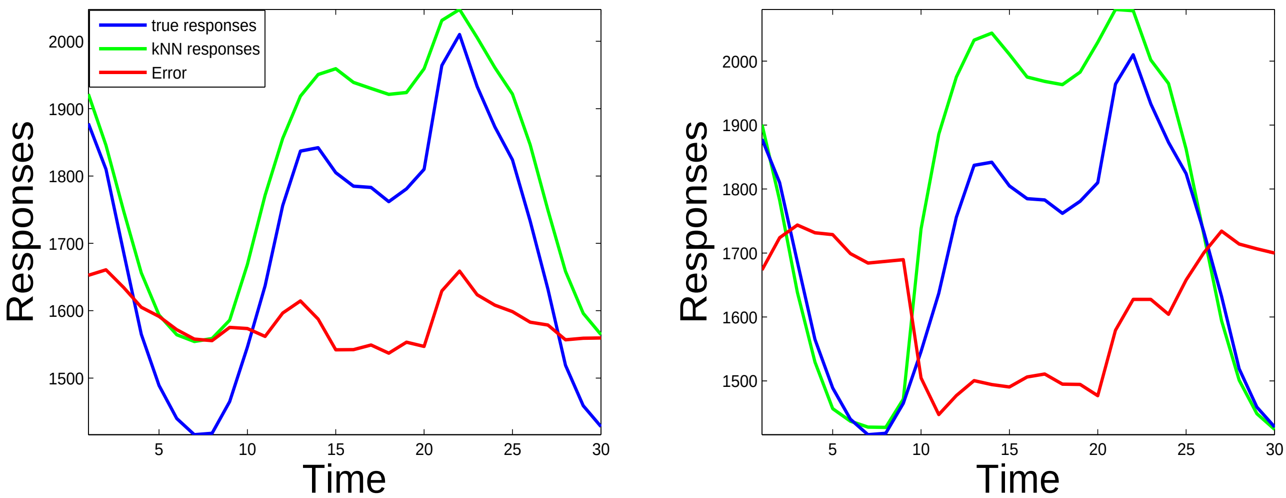

Рис. 3. Прогноз с использованием преобразования по OX (справа) и без (слева)

Видно, что уже на такой небольшой выборке заметно улучшение прогноза по сравнению с алгоритмом, не использующим преобразование по оси OX – в особенности на первых точках прогноза.

Рассмотрим теперь в 10 раз больший временной ряд (тот же самый временной ряд, но с большей историей) и получим следующие результаты:

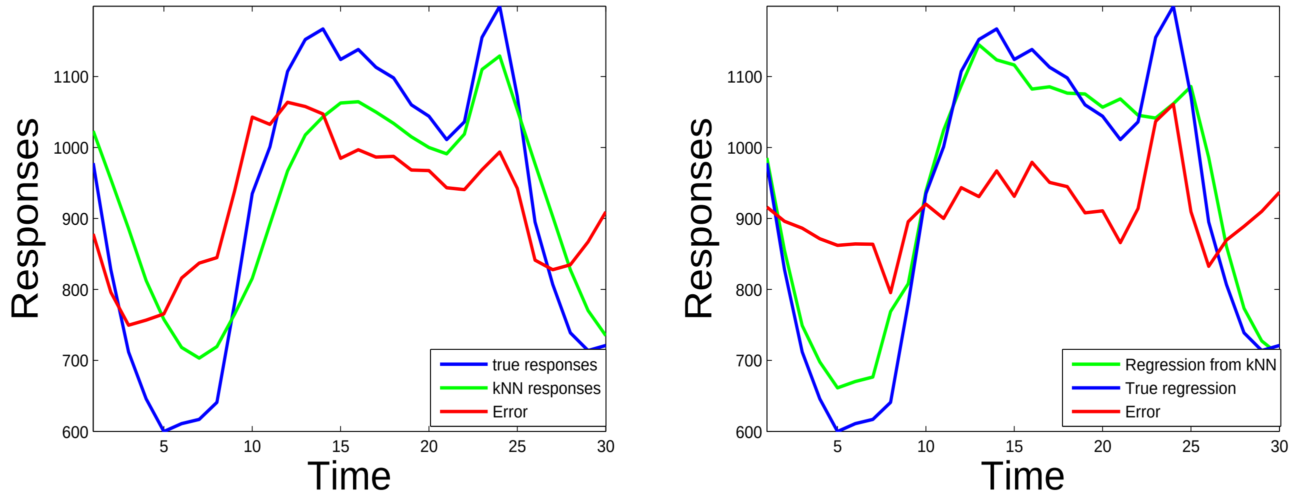

Рис. 4. Прогноз с использованием преобразования по OX (справа) и без (слева)

Алгоритм, использующий преобразование по оси OX прогнозирует намного лучше, нежели алгоритм, не использующий это преобразование. Кроме того, на такой большой выборке первый алгоритм затрачивает значительно меньшее время и ресурсов вычислительной машины, так как ищет коэффициенты преобразований не для всех возможных интервалов временного ряда, а для уже отобранных потенциальных соседей.

Точность прогноза сильно зависит от выбора значений параметров – это количество соседей k, длина предыстории L и параметр ε (максимальная ошибка при совмещении точек экстремумов потенциальных соседей и предыстории). Как уже было сказано во главе «Описание алгоритма», количество потенциальных соседей m пропорционально значению параметра ε. Алгоритм отбирает k наилучших из m потенциальных соседей после преобразований. Кажется, что если увеличивать параметр ε, то это не повлияет на результат – лучшие останутся лучшими и качество прогноза не изменится. Для опровержения этого был проведен следующий эксперимент, в котором исследовался функционал ошибки E прогноза от изменяющегося параметра ε . Результаты эксперимента представлены на следующем графике:

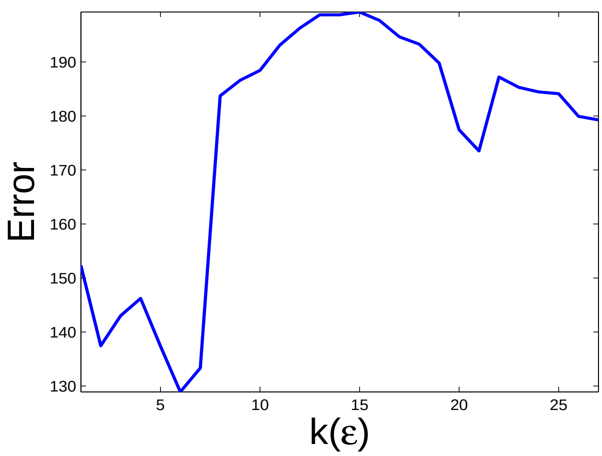

Рис. 5. Значение функционала ошибки от параметра ε

При увеличении параметра ε функционал ошибки E возрастает, что ещё раз подтверждает предположение о том, что преобразование по оси OX имеет большое значение.

Заключение

Итак, в работе был рассмотрен алгоритм локального прогнозирования, использующий афинные преобразования (сжатие, сдвиг) временного ряда по осям OX и OY и основанный на методе kNN (k ближайших соседей). Алгоритм был протестирован на модельных и реальных данных и показал хорошие результаты. В ходе сравнения работы алгоритма с алгоритмом, использующим преобразования только по оси OY, выяснилось, что преобразование по оси OX играет достаточно большую роль в прогнозировании методом kNN, не только улучшая качество прогноза, но и уменьшая затрачиваемые вычислительные ресурсы.

Литература

[1] McNames J.,*Innovations in local modeling for time series prediction* // Ph.D. Thesis, Stanford University, 1999.

[2] Журавлев, Ю. И., Рязанов, В. В., и Сенько, О. В. *Распознавание. Математические методы. Программная система. Практические применения.* // Фазис, Москва, 2005.

[3] Магнус, Я. Р., Катышев, П. К., Пересецкий, А. А. *Эконометрика* // Дело, 2004, стр. 34-37

[4] Федорова, В. П., *Локальные методы прогнозирования временных рядов* // Москва, 2009.

[5] Воронцов, К. В. Курс лекций *Математические методы обучения по прецедентам*

[6] Временные ряды прогнозирования электроэнергии *http://www.neural-forecasting-competition.com*

[7] Дуда, Р., Харт, П. *Распознавание образов и анализ сцен* // Мир, Москва, 1976