АЛГОРИТМ ПРОГНОЗИРОВАНИЯ СТОИМОСТИ АКЦИЙ РОССИЙСКИХ БАНКОВ

Баюк О.А., Берзин Д.В.

Финансовый университет при Правительстве Российской Федерации

Россия, г. Москва, Ленинградский проспект, д. 49

oabayuk@fa.ru, dberzin@fa.ru

Аннотация: В работе исследованы варианты применения некоторых известных методов прогнозирования временных рядов для прогнозирования стоимости акций российских банков. Авторами исследована зависимость точности прогноза значений стоимости акций от используемого алгоритма прогнозирования. Кроме того, предложено усовершенствование применяемых алгоритмов. Исследована эффективность применения исследуемых методов на известных данных для некоторых банков российской банковской системы.

Ключевые слова: банки, акции, прогноз, скользящее среднее, экспоненциальное сглаживание, экстраполяция многочленом.

Введение

Для прогнозирования временных рядов часто используются метод скользящего среднего и метод экспоненциального сглаживания. Мы исследовали эффективность использования указанных алгоритмов для прогноза значений стоимости акций российских банков. Для анализа были использованы данные по пяти российским банкам: Сбербанк России, ВТБ, РОСБАНК, МОСОБЛБАНК и Московский кредитный банк. Важным фактором, повлиявшим на отбор данных, является доступность этих данных по котировкам, так как не все брокерские сайты предоставляют информацию по банкам. Для исследования были использованы данные брокерского портала Finam.ru о ценах закрытия торгов на московской бирже.

Временной ряд – это случайная функция дискретного аргумента, как правило, моментов времени, интервалы, между которыми постоянны. Временной ряд задается как множество значений

$(y_1, y_2, \ldots, y_n)$, для которых известны значение момента времени, к которому относится Y1, и шаг по времени, а y_i – значения функции (зависимая переменная), i=1,..,n. Задача прогнозирования заключается в предсказании одного или нескольких следующих членов этого ряда.

1 Метод скользящего среднего

Метод скользящего среднего используются для сглаживания и прогнозирования временных рядов. Этот метод не столь эффективен в случаях, когда такая тенденция нарушается, например, при стихийных бедствиях, военных действиях, общественных беспорядках, при резком изменении параметров внутренней или внешней ситуации (уровня инфляции, цен на сырье); при коренном изменении плана деятельности фирмы, терпящей убытки. Основная идея метода скользящего среднего состоит в замене фактических уровней исследуемого временного ряда их средними значениями, при этом сглаживаются случайные колебания. Таким образом, в результате получается сглаженный ряд значений исследуемого параметра, позволяющий более четко выделить основную тенденцию его изменения.

2 Экспоненциальное сглаживание

Экспоненциальное сглаживание используется для выравнивания временных рядов. При этом для определения сглаженного уровня используются значения только предшествующих уровней ряда, взятые с определенными весами. Вес наблюдения уменьшается по мере удаления его от момента времени, для которого вычисляется сглаженное значение.

Пусть имеется временной ряд: $y_1, y_2, \ldots, y_n$. Соответствующие сглаженные значения уровней $S_1, S_2, \ldots, S_n$:

$$S_k = \alpha S_{k-1} + (1-\alpha) y_k = y_k + \alpha (S_{k-1} - y_k),$$

где α – параметр сглаживания $(0 < \alpha < 1)$, $(1-\alpha)$ – коэффициент дисконтирования. Обычно $(0{,}1 < \alpha < 0{,}3)$.

3 Экстраполяция приближенным выражением для многочлена Тейлора

Для вычисленя прогнозируемых значений цен акций предлагается использовать приближение полинома Тейлора

$$y(x+h) \approx y(x) + y'(x)h + \frac{1}{2} y''(x) h^2 + \frac{1}{6} y'''(x) h^3.$$

Производные в последней формуле заменяются разделенными разностями соответствующих порядков

$y'(x) \approx \frac{y_4 - y_3}{h}$,

$$y"(x) \approx \frac{y_4 - 2y_3 + y_2}{h^2},$$

$y'''(x) \approx \frac{y_4 - 3y_3 + 3y_2 - y_1}{h^3}$.

где

$y_1 = y(x_1), \quad y_{i+1} = y(x_1 + i \cdot h), i = 1, 2, 3.$

Таким образом, прогнозируемое значение цены акции в очередной момент времени определяется приближенной

$y_5 \approx \frac{1}{6}(5y_1 + 6y_2 - 15y_3 + 10y_4)$,

Проведенные расчеты показали, что использование этого алгоритма весьма эффективно.

Нами было исследовано применение алгоритмов скользящего среднего и экспоненциального сглаживания совместно с описанным выше алгоритмом экстраполяции для прогнозирования цен акций пяти российских банков. Для этого использовались исходные временного ряда с шагом один день. Результаты расчетов представлены в таблице 1.

Кроме того, была исследована эффективность работы алгоритмов для исходного временного ряда с шагом семь дней. Проведенные расчеты показали, что среднее значение ошибки вычисления экстраполируемого значения в этом случае составляет 10% истинного значения, максимальное значение ошибки вычисления экстраполируемого значения при экстраполяции на три шага, то есть на 21 день – 20% истинного значения.

Таблица 1

Название банка	**Сбербанк**	**ВТБ**	**Российский банк**	**Московский областной банк**	**Московский кредитный банк**
Интервал прогноза при использовании скользящего среднего	2 дня	2 дня	2 дня	2 дня	2 дня
Точность прогноза при использовании скользящего среднего	0,5 %	4 %	1 %	1 %	0,5 %
Интервал экстраполяции	6 дней	1 день	7 дней	4 дня	3 дня
Точность экстраполяции	5 %	5 %	3 %	5 %	2 %
Интервал прогноза при использовании экспоненциального сглаживания совместно с экстраполяцией	5 дней	1 день	3 дня	4 дня	6 дней
Точность прогноза при использовании экспоненциального сглаживания совместно с экстраполяцией	5 %	10 %	2 %	5 %	2 %

Заключение

Использование для прогнозирования стоимости акций российских банков целесообразно и обеспечивает приемлемую для практики точность. В данной работе описан выбор значения параметра в процедуре экспоненциального сглаживания. Предложенный нами алгоритм экстраполяции позволяет усовершенствовать метод прогнозирования.

Литература

1. *Козлов А.Ю., Мхитарян В.С., Шишов В.Ф.* Статистический анализ данных в MS Excel. Учебное пособие. – М.: ИНФРА-М, 2013. – 320 с.
2. *Литвинчук С.Ю.* Информационные технологии в экономике. Анализ и прогнозирование временных рядов с помощью EXCEL . Учебное пособие. – Нижний Новгород: ННГАСУ, 2010. – 78 с.
3. *Христиановский В.В., Щербина В.П.* Анализ временных рядов в экономике: практика применения. Учебное пособие. – Донецк: ДонНУ, 2011. – 125 с.

АЛГОРИТМ ОЦЕНКИ ВОСПРОИЗВОДЯЩЕГО ЯДРА ДЛЯ РЕШЕНИЯ ЗАДАЧИ МАШИННОГО ОБУЧЕНИЯ МЕТОДОМ КОЛЛОКАЦИИ

Бывшев В.А.

Финансовый университет при Правительстве Российской Федерации,
г. Москва, Ленинградский проспект, д.49
VByvshev@fa.ru

Аннотация: Обсуждается алгоритм оценки воспроизводящего ядра, необходимого для решения задачи машинного обучения методом коллокации. Метод коллокации восходит к Л.В. Канторовичу и состоит в восстановлении искомых функций по значениям функционалов. Алгоритм базируется на теореме Лоэва об интерпретации воспроизводящего ядра как автоковариационной функции и концепции Винера оценки автоковариационной функции стационарного ряда по индивидуальной реализации

Ключевые слова: Задача и метод машинного обучения, задача и метод коллокации, искомая функция и функционалы, воспроизводящее ядро.

Введение

Метод коллокации восходит к лауреату Нобелевской премии Л.В. Канторовичу [1] и состоит в восстановлении искомых функций по определённым на них значениям функционалов. В данной работе сначала излагаются задача и метод коллокации во взаимосвязи с постановкой задачи и методом машинного обучения. Затем строится алгоритм оценки воспроизводящего ядра, необходимого для решения методом коллокации задачи машинного обучения.

1 Задача машинного обучения

Задачи машинного обучения по прецедентам состоят в следующем [2], [3]. Задано множество X объектов $x \in X$ и задано множество Y допустимых ответов у $\in Y$. Каждый объект $x \in X$ описывается набором признаков $a(x) = (a_1(x), \dots, a_q(x))$, и с этим набором x отождествляется: $x \sim a(x)$. Каждый признак $a_j(x)$ имеет смысл некоторой характеристики объекта x и формально может интерпретироваться как отображение множества объектов X в множество A_j возможных значений признака a_j, то есть $a_j: X \to A_j$. В зависимости от типа множества A_j признак a_j может быть бинарным, категориальным или количественным. Отметим, что категориальные признаки можно квантифицировать, поэтому будем предполагать, что все признаки в наборе $a(x)$ являются количественными, то есть $a(x) \in R_q$.

По предположению, существует неизвестная функция $f: X \to Y$, значения которой $y_i = f(x_i)$ известны на конечном множестве объектов $(x_1, x_2, \dots, x_n) \subset X$. Пары $(a(x_i), y_i)$ именуются прецедентами; совокупность $S = (a(x_i), y_i)_{i=1}^{N}$ прецедентов образуют выборку. Выборку разделяют на две части: на обучающую выборку $S_l = (a(x_i), y_i)_{i=1}^{n}$ и контролирующую выборку $S_c = \left(a(x_{(i)}), y_{(i)}\right)_{(i)=1}^{(n)}$. Требуется по выборке S построить функцию $\hat{f}: X \to Y$, приближающую неизвестную функцию $f: X \to Y$ на всём множестве X.

Схема решения задачи машинного обучения состоит из трёх шагов [2].

Шаг 1. Задаётся некоторое параметрическое семейство Φ функций $\varphi(a; \theta)$:

(1.1) $$\Phi = \{\varphi(a; \theta) \big| \varphi: R_q \times \Theta \to \mathrm{Y}\}.$$