

Об одном методе прогнозирования динамики ценных бумаг

Мушег С. Саакян

Институт проблем информатики и автоматизации НАН РА
e-mail m.sahakian-alumni@isc.ac.uk

Аннотация

В статье приводится метод прогнозирования динамики ценных бумаг и результаты испытаний программной реализации этого метода.

1. Введение

В настоящее время существует много подходов к решению задачи предсказания (прогнозирования) временных рядов, соответствующих последовательным значениям валютного обменного курса. Среди подходов отметим изложенные в [1,2]. На динамику конкретной ценной бумаги влияют разнообразные факторы (это могут быть макро- и микроэкономические данные, торговые данные, оценки общественно-политических событий). В большинстве существующих моделей учитываются почти все эти разнообразные факторы. Это приводит к привязанности большинства моделей к какому-то определенному рынку.

Особенность этой работы заключается в том, что в построенной модели учитывается только курс ценной бумаги, что позволяет свести ее к более практически реализуемой математической задаче, которая универсальна для любого рынка. Ниже предлагается алгоритм нахождения оптимальной длины предыстории временного ряда, на основании которой производится прогнозирование, и формулируется метод такого прогнозирования. Модель реализована на языке СИ и является удобным приложением для принятия решений экспертом при осуществлении операций на Фондовом рынке. В качестве примера рассматривались ряды, составленные из значений курса доллара Центрального банка РА за N дней.

Учитывая специфику конкретной предметной области (курса доллара Центрального банка РА), исходную временную последовательность $B = \{b_1, b_2, \dots, b_N\}$ заменяем в модели на

последовательность $R = \{R_1, R_2, \dots, R_{N-1}\}$, где $R_i = \ln \left(\frac{b_{i+1}}{b_i} \right)$, $i \in (1, 2, \dots, N-1)$. В качестве

пороговых значений курса считаются выбранными два числа q_1 и q_2 . При этом, если $q_1 \leq R_i \leq q_2$, то курс на i -ый день считаем стабильным, если $R_i > q_2$, то курс на i -ый день возрос, если $R_i < q_1$, считаем что курс на i -ый день понизился. В соответствии с приведенными значениями изменений курса, сопоставив каждому i -ому элементу последовательности R элемент из алфавита $A = \{0, 1, -1\}$, получим слово $C = c_1 c_2 \dots c_{N-1}$ длины $N-1$ в алфавите A . Подберем константы q_1 и q_2 так, чтобы количество элементов 0, 1 и -1 в слове C было

"примерно одинаковым" т.е. отличалось на ± 1 . Наша цель — предложить метод, который позволит спрогнозировать следующую после c_{N-1} в слове C букву c_N .

Обозначим множество последовательных векторов вложения длины k вида

$$h_i = (c_i, c_{i+1}, \dots, c_{i+k-1}), \quad 1 \leq i \leq N-k+1,$$

для исходного ряда B . Эти векторы называются k -векторами вложения и образуют траекторную матрицу ряда B [3]

$$H_B = \begin{pmatrix} c_1 & c_2 & \dots & c_{N-k+1} \\ c_2 & c_3 & \dots & c_k \\ \vdots & \vdots & \ddots & \vdots \\ c_k & c_{k+1} & \dots & c_N \end{pmatrix}.$$

Далее используется термин окно вместо вектора вложения длины k и предполагаемый метод состоит из трех последовательных шагов: нахождение минимального окна, прогнозирование по заданному окну и самообучение.

2. Нахождение минимального окна

Множество окон длины k для последовательности C разобьем на, вообще говоря, пересекающиеся, подмножества M_1^k, M_0^k, M_{-1}^k следующим образом. Если r -элемент в слове C , следующий за окном h длины k , то h включается в подмножество (набор) M_r^k .

Проделаем вышеуказанную процедуру по $k \in (1, 2, \dots, N-1)$. Получим $N-1$ троек наборов (разных длин k), из которых выберем одну, которая удовлетворяет следующим условиям:

1. $\frac{|M_{i,j}|}{N-k-1} < 0,05$, где $M_{i,j} = M_i^k \cap M_j^k$, при $i, j \in A$.
2. длина наборов $k_{\min}$ минимальная, среди таких троек, которые удовлетворяют условию 1, окно длины $k_{\min}$ назовем минимальным

3. Алгоритм прогнозирования по выбранному окну

Ниже предлагается алгоритм прогнозирования последующего элемента.

Далее, для упрощения обозначений опустим индекс k в M_i^k . Построим множества M_1, M_0, M_{-1} для заранее выбранного окна по алгоритму, предложенному в п. 2 и договоримся не включать в указанные подмножества окно, которое в слове C стоит перед элементом c_N .

Определим расстояние p_i окна $h = (h_1, h_2, \dots, h_k)$ от множества M_i следующим образом:

$$p_i(h) = \min_{m^i \in M_i} \left(\frac{\sqrt{\sum_{j=1}^k (h_j - m_j^i)^2}}{\frac{1}{|M_i|} \sum_{m^i \in M_i} \sqrt{\sum_{j=1}^k (h_j - m_j^i)^2}} \right).$$

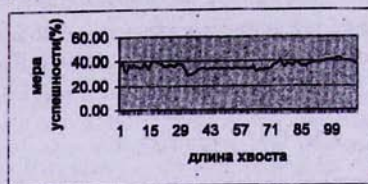
Вычислим расстояния p_{-1}, p_0, p_1 окна h , которое в слове S стоит перед элементом c_N , от множеств M_{-1}, M_0, M_1 . Заметим, что если окно $h \in M_1$, то расстояние $p_1(h) = 0$. Прогнозируемый элемент c_N будем считать равным 1, если расстояние $p_1(h)$ минимальное среди всех $l \in (-1, 0, 1)$.

4. Самообучение

Для каждого окна длины k из R , где $k \in [k_{\min}, \dots, k_{\max}]$, перебираются все элементы, перед которыми есть окно такой длины, причем каждый такой элемент считается неизвестным и к нему применяется Алгоритм прогнозирования по выбранному окну. Реальное значение элемента сравнивается с прогнозируемым, и при положительном результате проверки прогнозирования считается успешным. Число успешных результатов прогнозирования деленное на длину выборки N дает меру успешности прогнозирования для окна данной длины. Оптимальным считается то окно, мера успешности которого наибольшая. Данные, полученные на основании этого окна, учитываются при принятии решения экспертом.

5. Результаты эксперимента

Произведен эксперимент на основании данных Центрального банка РФ по курсу доллара за период от 1 апреля 2000 г. до 16 апреля 2001г. В нижеприведенном графике показана зависимость меры успешности от длины окна.



Полученная оптимальная длина окна, на основе которой произведено прогнозирование, равна 102, мера успешности составила 42.2%. Спрогнозированный элемент совпал с реальным значением.

Было произведено 68 случайных экспериментов на различных интервалах имеющегося курса. Мера успешности не превосходила 50%, в то время как прогнозируемые значения совпали с реальными в 56 случаях. Оптимальная длина окна зависела от длины рассматриваемой последовательности. Экспериментально она не превысила половины длины рассматриваемой последовательности.

Литература

1. C. Lee Giles, B. G. Horne and T. Lin. Learning a class of large finite state machines with a recurrent neural network. *Neural Networks* 8(9), 1359-1365, 1995.
2. Кокс Д. Р., Окс Д. Анализ данных типа времени жизни. М. Финансы и статистика. 1988.
3. Голяндина Н. Э. Матод "Гусеница" — SSA: анализ временных рядов. Учебное пособие СПб: Изд-во С-Петербургского университета. 2004. 78 с.

Արժեթղթերի դիմամիկայի կանխագուշակման մի մեթոդի մասին

Մ. Ս. Սահակյան

Ամփոփում

Հոդվածում ներկայացված են արժեթղթերի դիմամիկայի կանխագուշակման մի մեթոդ և դրա ծրագրային իրագործման փորձարկումների արդյունքները: