

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»**

Кафедра

Математической экономики

Прогнозирование цен на нефть на основе чебышевской задачи

наилучшего приближения

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 441 группы

направление 09.03.03— Прикладная информатика

механико-математического факультета

Суркова Ярослава Алексеевича

Научный руководитель

зав. каф., д.ф.-м.н., профессор

С.И. Дудов

Зав. кафедрой

зав. каф., д.ф.-м.н., профессор

С.И. Дудов

Саратов 2019

Введение. Все трейдеры в своей работе постоянно использует различные инструменты для технического анализа, позволяющие повысить эффективность сделок на рынке и увеличить вероятность успеха. Одними из таких инструментов являются технические индикаторы, полученные путем осуществления ряда математических действий.

Данная тема является актуальной, так как торги на бирже являются неотъемлемой частью мировой экономики. А для совершения успешных сделок трейдерам необходимы более совершенные инструменты технического анализа.

Цели, преследуемые в ходе выполнения работы:

- Проанализирован рынок нефти;
- Проанализированы графики, по которым работают трейдеры;
- Построен индикатор прогнозирования на основе полиномиальной интерполяции;
- Проведён вычислительный эксперимент с целью тестирования индикатора.

Основная часть. Постановка задачи чебышевского приближения. Пусть функция $f(x)$ задана таблицей своих значений $y_k = f(x_k)$, $k = 0, \dots, N$, где $N \geq n + 1$.

x	x_0	x_1	x_2	\dots	x_N
y	y_0	y_1	y_2	\dots	y_N

(1)

Где $x_0 < x_1 < x_2 < \dots < x_N$. Всякий алгебраический полином $P_n(A, x) = \sum_{i=1}^n a_i x^i$, по отношению к заданной выше таблице, имеет максимальное уклонение:

$$\phi(A) = \max_{0 \leq k \leq N} |y_k - P_n(A, x_k)|.$$

Пусть $A^* \in R^{n+1}$ такой вектор для которого

$$\rho = \phi(A^*) = \min_{A \in R^{n+1}} \phi(A).$$

Тогда полином $P_n(A^*, x)$ называется полиномом лучшего приближения для таблицы (1).

Решение задачи наилучшего приближения (построение полинома $P_n(A^*, x)$) существенно зависит от соотношения между значениями порядка полинома n и «длины» таблицы N . Рассмотрим случай, когда $N = n + 1$.

Теорема 1. Для того, чтобы полином $P_n(A, x)$ был полиномом наилучшего приближения для таблицы (1) при $N = n + 1$ необходимо и достаточно, чтобы при некотором $h \neq 0$ выполнялось условие:

$$P_n(A, x_k) - y_k = (-1)^k \cdot h, \text{ где } k = 0, \dots, n + 1,$$

что так же гарантирует выполнение условия $|h| = \rho$.

Доказательство. Вопрос о существование полинома вида $P(A, x)$ сводится к решению следующей системы уравнений:

Последнее доказывает существование у функции $n + 1$ нуля. Но полином n -го порядка, имеющий на конечном интервале более n нулей, должен быть тождественно равным нулю. Таким образом:

$$Q_n(x) \equiv 0 \text{ и } P_n(A_1, x) \equiv P_n(A^*, x).$$

Следовательно, предположение $|h| > \varrho$ не верно и теорема полностью доказана.

Сведение дискретной чебышевской задачи к задаче линейного программирования. Теперь рассмотрим случай, когда в таблице (1) $N > n + 1$.

Приведем критерий решения этой задачи, полученный в (3). Для этого введем дополнительные обозначения

$$\begin{aligned} R_1(A) &= \{x_i \in \{x_k\}, k \in [0 : N] : \phi(A) = P_n(A, x_i) - y_i\}, \\ R_2(A) &= \{x_i \in \{x_k\}, k \in [0 : N] : \phi(A) = y_i - P_n(A, x_i)\}. \end{aligned}$$

Теорема 2. Для того, чтобы полином $P_n(A^*, x)$ был полиномом наилучшего приближения для таблицы (1) при $N > n + 1$, необходимо и достаточно, чтобы существовал упорядоченный набор точек $\{x_{i_k}\}, k = [1 : n + 2], x_{i_1} < x_{i_2} < \dots < x_{i_{n+2}}$ такой, что, если

$$x_{i_k} \in R_1(A^*)(R_2(A^*)), \text{ то } x_{i_{k+1}} \in R_2(A^*)(R_1(A^*)), i = \overline{1, n + 1}.$$

Введем обозначения:

$$\begin{aligned} B_i &= (1, x_i, x_i^2, \dots, x_i^n), b_i = -y_i, B_{N+i+1} = -(1, x_i, x_i^2, \dots, x_i^n), b_{N+i+1} = y_i, \\ & i \in [0 : N]. \end{aligned}$$

Тогда, учитывая, что полином $P_n(A, x)$ можно представить в виде скалярного произведения $P_n(A, x) = \langle A(1, x, x^2, \dots, x^n) \rangle$, а задачу $\phi(A) = \max_{i \in [0 : N]} |P_n(A, x_i) - y_i| \rightarrow \min_{A \in R^{n+1}}$ можно записать в виде:

$$\max_{i \in [0 : 2N]} \{\langle B_i, A \rangle + b_i\} \rightarrow \min_{A \in R^{n+1}}. \quad (4)$$

Справедлива следующая теорема.

Теорема 3. Задача (4) эквивалентна задаче линейного программирования следующего вида (относительно неизвестных $A \in R^{n+1}$ и a_{n+1}):

$$\begin{cases} a_{n+1} \rightarrow \min \\ a_{n+1} - \langle B_i, A \rangle - b_i \geq 0, i \in [1 : 2N]. \end{cases} \quad (5)$$

При этом, если (a_{n+1}^*, A^*) – решение задачи (5), то A^* – решение задачи (4) и $a_{n+1}^* = \max_{i \in [0:N]} |P_n(A^*, x_i) - y_i|$.

Доказательство. 1) Пусть $(a_0^*, a_1^*, \dots, a_n^*, a_{n+1}^*) = (A^*, a_{n+1}^*)$ является решением задачи (5). Тогда $a_{n+1}^* \geq \langle B_i, A^* \rangle + b_i$ для всех значений индекса i из $[1 : 2N]$.

Следовательно,

$$a_{n+1}^* \geq \max_{i \in [1:2N]} \{\langle B_i, A^* \rangle + b_i\} \geq \min_{A \in R^{n+1}} \max_{i \in [1:2N]} \{\langle B_i, A \rangle + b_i\} = L. \quad (6)$$

С другой стороны, если точка $A^0 = (a_0^0, a_1^0, a_2^0, \dots, a_n^0) \in R^{n+1}$ является решением задачи (4), то есть

$$\max_{i \in [1:2N]} \{\langle B_i, A^0 \rangle + b_i\} = \min_{A \in R^{n+1}} \max_{i \in [1:2N]} \{\langle B_i, A \rangle + b_i\} = L,$$

то $\langle B_i, A^0 \rangle + b_i \leq L, i \in [1 : 2N]$, а значит, точка $(A^0, L) \in R^{n+2}$ удовлетворяет условиям задачи (5), не превышает значение L .

$$a_{n+1}^* \leq L. \quad (7)$$

Таким образом, из (6) и (7) следует

$$a_{n+1}^* = L. \quad (8)$$

Далее из (6) и (8) получаем

$$\max_{i \in [1:2N]} \{\langle B_i, A^* \rangle + b_i\} = \min_{A \in R^{n+1}} \max_{i \in [1:2N]} \{\langle B_i, A \rangle + b_i\},$$

что означает, что A^* решение задачи (4).

2) Пусть теперь $A^0 \in R^{n+1}$ – решение задачи, то есть имеем

$$\max_{i \in [1:2N]} \{ \langle B_i, A^0 \rangle + b_i \} = \min_{A \in R^{n+1}} \max_{i \in [1:2N]} \{ \langle B_i, A^* \rangle + b_i \} = L,$$

Отсюда получим

$$L \geq \langle B_i, \vec{a}^0 \rangle + b_i, i \in [1 : 2N].$$

Что означает, что точка $(A^0, L) \in R^{n+2}$ удовлетворяет условиям задачи (5). Поэтому, если a_{n+1}^* является минимальным значением, для которого выполняются условия задачи (5), то

$$a_{n+1}^* \leq L.$$

А с другой стороны, как уже было показано в пункте 1), для этого значения справедливо

$$a_{n+1}^* \geq L.$$

Таким образом

$$a_{n+1}^* = L.$$

Итак, показано, что точка $(A^0, L) \in R^{n+2}$ удовлетворяет условиям задачи (5) и при этом, L - минимальное из возможных значений для a_{n+1}^* . Следовательно, эта точка $(A^0, L) \in R^{n+2}$ является решением задачи (5).

3) Таким образом, показано, что имея решение задачи (5) в виде (A^*, a_{n+1}^*) получается решение задачи (4) в виде вектора A^* .

А если имеем решение задачи (4) – вектора A^* , то получаем в виде точки $(A^*, L) \in R^{n+2}$, где L определяется из (6), решение задачи (5). Это говорит об эквивалентности задач. Теорема доказана.

Схема бэк-тестирования индикатора, использующего чебышевскую задачу наилучшего приближения. Пусть исторические значения цен некоторого актива (ценных бумаг) записаны в таблице.

x	x_0	x_1	x_2	\dots	x_N
y	y_0	y_1	y_2	\dots	y_N

где y_N цена закрытия в период времени x_i .

для всех $x_k : x_0 < x_1 < \dots < x_N$.

1) Выбираем степень алгебраического полинома n .

2) Выбираем количество используемых узлов $m \geq n + 2$. Для подсчета значений индикатора в точках $\{x_k\}$ с целью его последующего тестирования выполняются следующие действия:

3) Полагаем $i = 0$.

4) Решается задача

$$f(A) = \max_{k \in [i:m+i+1]} |y_k - P_n(A, x_k)| \rightarrow \min_{A \in R^{n+1}}.$$

5) Пусть вектор коэффициентов A_i является решением задачи. В качестве значения индикатора $I_{m,n}(x)$ в точке x_{i+m+1} берётся

$$I_{m,n}(x_{i+m+1}) = P_n(A_i, x_{i+m+1}).$$

Если $i + m + 1 < N$, то полагаем, $i := i + 1$, передаём управление в пункт 4. Иначе, если $i + m + 1 = N$, расчёт окончен. После проделанной работы получаем значения индикатора $I_{m,n}(x)$ для x_{m+1}, \dots, x_n .

В итоге получим прогнозные значения цен в точках $x_{m+1}, x_{m+2}, \dots, x_N$.

Эти данные подлежат дальнейшему анализу для получения выводов об эффективности данного индикатора на фоне сравнения его значений с историческими значениями $y_{i+m+1}, i \in [0 : N - m - 1]$ соответственно, а так же с эффективностью прогноза других индикаторов.

Тестирование индикатора на примере цен закрытия нефтяного рынка. План проведения экспериментов. Считаем, что данные о котировках цен на нефть заданы в виде

$$y_k(x_k), k \in [1 : N],$$

где $y_k(x_k)$ значение цены закрытия фьючерсов на нефть.

- 1) Задаем степень полинома $n < N$, $m \geq n + 2$
- 2) Подсчитываем значение индикатора на основе решения задачи чебышевского приближения:

$$I_{m,n}(x_{i+m+1}) = P_n(A_i, x_{i+m+1}), \text{ где } P_n(A_i, x_{i+m+1}) = \sum_{i=1}^n a_i x^i.$$

- 3) Проводим серию экспериментов для $n = 2$, $m = 3$.
- 4) Анализируем полученные результаты.

Правила торговли с использованием индикатора. Для данного индикатора актуальны те же правила торговли как и для скользящих средних. Если линия тренда направляется вверх, то прогнозируется рост на рынке, если вниз, то прогнозируется падение.

К сигналам можно отнести следующую особенность, если ценовой график растёт вместе с линией тренда, то прогнозируется рост и можно ставить заявку на покупку, если ценовой график вместе с линией тренда направляется вниз, то прогнозируется падение и ставится заявка на продажу. Так же с помощью данного индикатора можно осуществлять торговлю на пересечениях с ценовым графиком. Если ценовой график пробивает линию тренда сверху вниз, то ставится заявка на покупку, если снизу вверх, то ставится заявка на продажу.

Результаты вычислительных экспериментов. Проведем вычислительные эксперименты, используя разработанную программу, реализующую получение значений индикаторов на основе решения задачи по интерполированию. Задаем таблицы значений $y_k = y(x_k)$, $k \in 1 : N$, которая содержит в себе максимальные и минимальные значения котировок на нефть марки Brent за период с 01.01.2019 по 10.04.2019.

Дискретный график имеет вид:

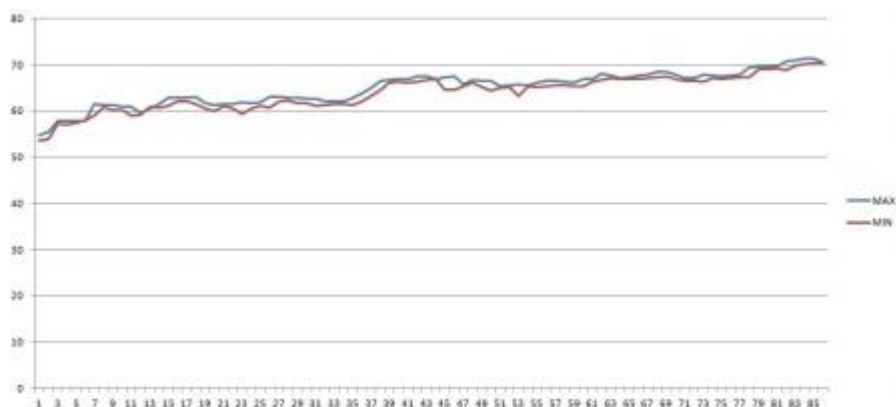


Рисунок 1—Значения максимальных и минимальных цен период с 01.01.2019 по 10.04.2019

Приведем пример расчетов по используемой программной разработке для первых трех прогнозных значений индикатора.

Вводим $n = 2$ и $m = 3$: Получаем:

$$I_{3,2}(x_{12}) = 65, 50$$

$$I_{3,2}(x_{13}) = 65, 50$$

$$I_{3,2}(x_{14}) = 65, 38$$

Проведение вычислительного эксперимента. Для проведения эксперимента потребуются значения максимальных и минимальных котировок цен. Для этого сохраняем данные минимальных значений цен в файл «min.txt» и максимальных значений цен в файл «max.txt». На выходе получаем график индикатора с прогнозными значениями.

Сравним наш индикатор с другим, подобным индикатором, уже давно зарекомендовавшим себя на рынке. Таким является SMA (простая скользящая средняя).

Данный индикатор схож с SMA в том, что он основан на истории и использует так же максимальные и минимальные значения цены.

Для проведения эксперимента будут использованы цены закрытия и сравним данный индикатор по количеству сигналов для совершения сделок с SMA. Для обоих индикаторов используется одна котировка нефти марки Brent с таймфреймом равным одному дню.

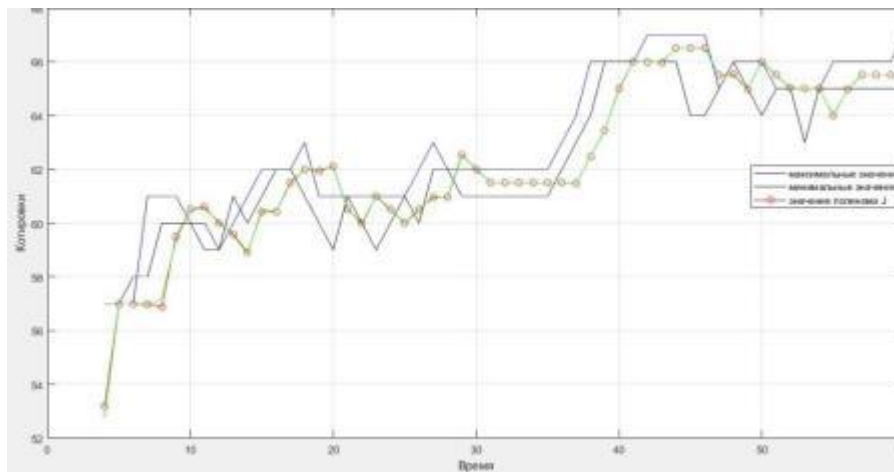


Рисунок 2— Значения максимальных и минимальных цен совместно с индикатором I

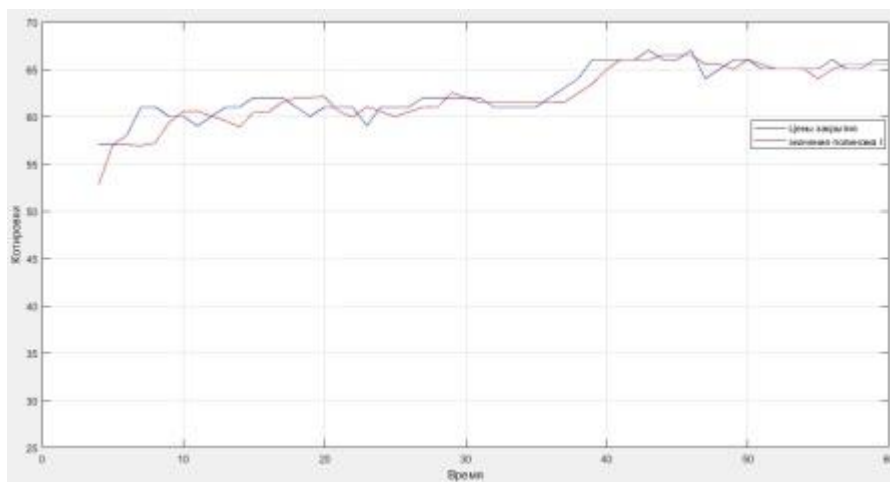


Рисунок 3— Значения цен закрытия совместно с индикатором I

Индикатор на основе Чебышевского приближения показал, что можно совершить 16 сделок, из которых больше половины принесут прибыль.

Индикатор SMA показал 26 сигналов и прибыльными оказались больше половины сделок. Данный индикатор один из наиболее часто используемых в техническом анализе.

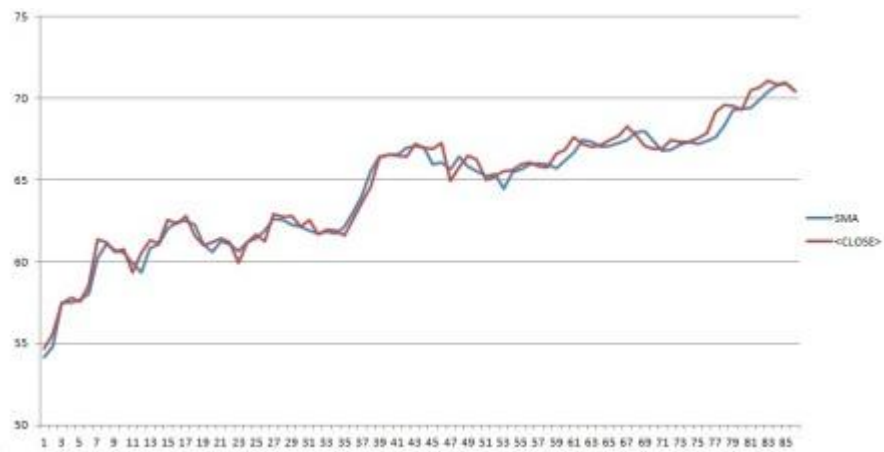


Рисунок 4— Значения цен закрытия совместно с индикатором SMA

В результате проведённого эксперимента были выявлены следующие преимущества и недостатки

Преимущества:

- Легко различимые сигналы.
- Даёт достаточное количество информации для анализа движения цены

Недостатки:

- Запаздывание
- Много ложных сигналов
- Недополучение прибыли.

Заключение. Прогнозирование означает специальные научные исследования конкретных перспектив развития какого-либо явления и находится во взаимосвязи с планированием, программированием и управлением.

Прогнозирование для отдельного объекта исследования или отдельного предприятия означает процесс формирования прогнозов его развития на основе предыдущих тенденций развития, изменяющихся во времени, т.е. показателей рядов динамики.

Исследование тенденции движения тренда цен, безусловно, важная и полезная часть прогнозирования. В работе был построен алгоритм получения будущих результатов от движения тренда цен. Все поставленные ранее цели и задачи были выполнены.