

УДК 336.764

В.В. Крюкова, П.А. Крюков

СТАТИСТИЧЕСКОЕ ПРОГНОЗИРОВАНИЕ ВАЛЮТНОГО КУРСА

Введение

Большинство торговых стратегий инвесторов основаны на отслеживании и идентификации текущего тренда и момента разворота цены инструмента. Под прогнозированием валютного курса в данной работе понимаем идентификацию текущего динамического поведения цены инструмента (направление и смену тренда).

Для прогнозирования валютных курсов используются методы технического анализа (ТА). «Технический анализ – это искусство распознавания на ранней стадии тенденции изменения цен и сохранения инвестиционной позиции до тех пор, пока не будет доказано, что тенденция делает свой поворот в обратном направлении» [1].

Методы ТА основаны на ряде аксиом, хорошо изложенных в работе А. Эрлиха [2]. Основные постулаты ТА определяют принципиальную возможность прогнозирования валютного курса, что в свою очередь предполагает возможность получения высоких прибылей от спекулятивных операций на валютном рынке.

Все методы ТА делят на три группы: графический, теория циклов и математико - статистический.

Основой графического метода является использование для прогнозирования рынка типовых графических фигур (линий поддержки/ сопротивления, фигур разворота и продолжения тенденции). Серьезным недостатком графического метода является субъективизм в трактовке подаваемых сигналов. Как отмечает Дж. Мэрфи [3], графический анализ достаточно субъективен, а результаты его оказываются часто противоречивыми.

В отличие от графического метода теория циклов позволяет определить не только сигналы изменения или продолжения тренда, но и момент времени реализации этого сигнала. Интерпретация цены как циклической структуры базируется на принципах: гармоничности, синхронности, пропорциональности и суммирования.

Примером теории циклов является волновая теория Эллиотта [4]. Достоинством метода является синтез графического и аналитического подходов. Недостатком – неспособность прогнозирования краткосрочных тенденций и «плохая формализованность» графических методов в целом.

Математико – статистический метод основан на анализе различных показателей, полученных в результате статистических расчетов по значениям цены инструмента (цены открытия / закрытия, максимальной, минимальной и т.д.). При достижении этими показателями некоторых пороговых значений появляется сигнал о смене текущей тен-

денции на рынке. К основным инструментам математико – статистического метода относят скользящие средние (СС), индикаторы и осцилляторы. Сигналы, подаваемые СС: если линия СС ниже кривой цены на графике, то текущий тренд - восходящий; если выше – нисходящий; их пересечение предупреждает о развороте тенденции.

Индикаторы тенденций позволяют выявлять участки тренда на графике, используются для подтверждения текущего тренда.

Осциллятор – модель колебаний, хорошо работает во флэте. Значение осциллятора рассчитывается по определенной формуле и показывает соотношение различных значений цен за некоторый интервал времени в процентах или абсолютных величинах. Для каждого осциллятора определяются границы (уровни) перепроданности / перекупленности рынка. Сигнал к продаже – значение осциллятора пересекает границу перекупленности, к покупке – границу перепроданности. Важным сигналом является расхождение (дивергенция) между направлениями движений цены и осциллятора, предупреждает о развороте тренда [5].

Для прогнозирования валютного курса в практической деятельности трейдера широко применяются все три метода технического анализа. Общим недостатком методов можно указать «ручной» способ и субъективизм принятия решения в реализации торговых операций. Правила торговли, найденные эмпирически, не имеют строго обоснования.

С развитием вычислительной техники и средств коммуникаций, появлением сети Интернет появилась возможность автоматизировать процесс принятия решения в виде механической торговой системы трейдера, позволяющей формализовать правила торговли, научно обосновать элементы принятой торговой стратегии.

Цель исследования: разработать значимую в статистическом смысле эконометрическую модель индикатора вероятности валютного курса, пригодную для прогнозирования направления движения цены инструмента (восходящий / нисходящий тренд) на валютном рынке FoRex.

Математическая постановка задачи

Пусть есть некоторое множество наблюдений (ситуаций, состояний рынка), относящихся к ρ ($\rho=2$) различным классам. Найти правило, которое с минимальным числом ошибок позволяет классифицировать новые наблюдения на основе предыдущих наблюдений и их классификаций. Пусть наблюдение описывается вектором X , а его классификация числом y ($y=0;1$). Имея последовательность вида $\{x_i^k, y_i\}$ из n наблюдений k переменных

и их классификаций y найти решающее правило классификации новых наблюдений.

Важно подчеркнуть, модель основана на оценке и изучении исторических данных со стороны влияния различных переменных (факторов – вектор X), определяющих тип процесса – ситуацию, состояние рынка (объект классификации).

Сформулируем постановку задачи в терминах теории восстановления зависимостей. Множество объектов – ситуаций X в соответствии с некоторой вероятностной мерой, плотность которой $P(x)$ неизвестна, появляются случайно и независимо. Количественные различия параметров объектов (ситуаций) позволяют отнести их к качественно различным классам с помощью некоторого правила, которое определяется условной вероятностью $P(y|x)$. Каждому вектору X ставится в соответствие число y , полученное в результате реализации случайного испытания согласно плотности $P(y|x)$. Переменная y принимает одно из двух значений, экономический смысл которых: 1 – рост, 0 – падение курса. Требуется по случайной независимой выборке объема n пар $\{x_i^k, y_i\}$ восстановить регрессию (функцию условного математического ожидания), то есть в классе функций $\{P_f(L(x, \alpha))\}$ найти функцию, наиболее близкую к регрессии $y(x) = \int yP(y|x)dy$.

В общем случае, задача прогнозирования валютного курса решается путем построения нелинейной динамической системы, восстанавливающей функцию условного математического ожидания, описывающую стохастическую зависимость валютного курса вида $\eta = L(g(X)) + \phi$, где $g(X)$ – вектор – функция предварительного преобразования переменных. Качество прогноза оценивается с помощью функции – индикатора вероятности вида:

$P(y_i=1) = P_f(L(x_i))$, $P(y_i=0) = 1 - P_f(L(x_i))$, то есть

$$P_f(L(X, \alpha)) = \begin{cases} \text{класс } 1, \text{ если } L(X, \alpha) > \psi \\ \text{класс } 0, \text{ если } L(X, \alpha) < \psi \end{cases},$$

где ψ – некоторый порог отсечения, α – параметр, принадлежащий множеству A . $P_f(L(X, \alpha))$ – монотонная функция распределения, область значений которой – интервал от 0 до 1.

В качестве индикаторной функции выбрана функция логистического распределения

$$P_f(L(x, \alpha)) = \frac{1}{1 + e^{-L(x, \alpha)}}.$$

$L(X, \alpha)$ – линейный предиктор,

$$L(X, \alpha) = b + \sum_{j=1}^k \alpha_j x_j,$$

где $x_j \in X$ ($j = 1, \dots, k$) – переменные (факторы), однозначно определяющие состояние рынка (системы) в момент времени t .

Для разработки и идентификации модели ин-

дикатора используется логит – регрессия. Выбор функции распределения является эмпирическим. Применение интеграла вероятности в качестве преобразования некоторой линейной функции позволяет получить результат в интервале от 0 до 1. Логистическая функция распределена в нужном интервале, находит вероятность дискретного события, экономический смысл которого: вероятность близка к 1 – текущий тренд восходящий, близка к 0 – тренд нисходящий. Она не линейна, точка перегиба кривой может описывать разворот тенденции, проста в применении и хорошо описана в литературе, поэтому выбрана для решения поставленной задачи. Кроме того, логит – модель допускает более широкие возможности в плане использования средств анализа качества полученной модели.

В логит - регрессии предполагается, что имеется функциональная зависимость (функция связи ϕ) между ожидаемым значением переменной EY и линейным предиктором $L(X, \alpha)$: $\eta = \phi(EY)$. Функция связи: $\log_e \{EY / (1 - EY)\} = \eta$.

Название «логит» происходит от названия простого способа сведения модели к линейной с помощью логистического (логит -) преобразования. Ожидаемое значение переменной EY – это вероятность. Рассуждая о переменной Y в терминах вероятности p , лежащей между 0 и 1, можно преобразовать эту вероятность по правилу: $p' = \log_e \{p / (1 - p)\}$.

Если произвести логистическое преобразование обеих частей равенства

$$y = \frac{1}{1 + e^{-L(x, \alpha)}},$$

получим стандартную модель линейной регрессии $p' = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_k x_k$.

Функция логистического распределения (логиста) равна

$$F(X) = \frac{1}{1 + e^{-X}}.$$

Величина $L(X, \alpha)$, определяемая значением функции $F(X)$ называется логитом $F(X)$. Модель с бинарной зависимой переменной с логистически распределенным отклонением называют логит – моделью [6, с. 337].

Оценка логит - регрессии

Логит-модели оценивают методом максимального правдоподобия (ММП) [7].

Показателями ее качества, являются информационные критерии Акайке (АС) и Шварца (SW):

$AC = -2 \ln(L_i) + 2k$ и $SW = -2 \ln(L_i) + k \ln(n)$, где k – количество регрессоров и n – количество наблюдений. Информационные критерии оценивают качество модели как баланс качества приближения к данным и количеством предикторов – сложностью модели, поэтому статистика состоит из штрафа за недостаточную подгонку к данным и

за «лишние» параметры. Оптимальная в смысле указанных критериев модель будет иметь минимальное значение критерия [7].

Вне зависимости от рассматриваемой модели, всегда можно оценить *объясненную долю дисперсии* переменной y в регрессионной модели, которая эквивалентна значению R^2 – *коэффициенту детерминации*. Даже если распределение зависимой переменной не является нормальным, эта величина помогает оценить, насколько хорошо подобранная модель согласуется с исходными данными (чем ближе к 1, тем лучше) [7].

Для моделей с бинарной зависимой переменной используется аналог коэффициента детерминации — псевдо- R^2 , который имеет максимальное значение много меньше единицы [6].

Для оценки качества модели логит - регрессии используется инструмент *ROC-анализ*, результат которого представим в виде *ROC - кривой (Receiver Operator Characteristic)*. Она показывает зависимость количества верно классифицированных положительных случаев от количества неверно классифицированных отрицательных случаев. В терминологии *ROC - анализа* первые называются истинно положительным, вторые – ложно отрицательным множеством. При этом предполагается, что у классификатора имеется некоторый параметр, варьируя который можно получать то или иное разбиение на два класса. Этот параметр называют порогом отсечения [8].

Ценность любого бинарного классификатора определяется двумя свойствами модели: чувствительностью и специфичностью.

Чувствительность модели S_e – это доля истинно положительных случаев, которые были правильно идентифицированы моделью

$$S_e = tp_R = \frac{tp}{tp + fn} \cdot 100\%,$$

где tp - верно классифицированные положительные случаи (истинно положительные); fn - положительные случаи, классифицированные как отрицательные (ложно отрицательные).

Специфичность модели S_p – это доля истинно отрицательных случаев, которые были правильно идентифицированы моделью

$$S_p = \frac{tn}{tn + fp} \cdot 100\%,$$

где fp - отрицательные случаи, классифицированные как положительные (ложно положительные); tn - верно классифицированные отрицательные случаи (истинно отрицательные).

Модель с высокой чувствительностью дает истинный результат при наличии положительного исхода (обнаруживает положительные случаи). Модель с высокой специфичностью дает истинный результат при наличии отрицательного исхода (обнаруживает отрицательные случаи).

Чем ближе построенная ROC - кривая к верхнему левому углу, тем выше предсказательная

способность модели. Невозможно одновременно повысить чувствительность и специфичность модели. Компромисс находится с помощью порога отсечения, так как он влияет на соотношение S_e и S_p . Порог отсечения нужен для применения модели на практике: относить новые случаи к одному из двух классов. Для определения оптимального порога нужно задать критерий. В практическом применении используют критерии:

1. Требование минимальной величины чувствительности (специфичности) модели. Например, нужно обеспечить чувствительность теста не менее 90%. В этом случае оптимальным порогом будет максимальная специфичность (чувствительность), которая достигается при 90% чувствительности (специфичности).

2. Требование максимальной суммарной чувствительности и специфичности модели
$$\psi = \max_k (Se_k + Sp_k)$$

3. Требование баланса между чувствительностью и специфичностью ($S_e \approx S_p$)
$$\psi = \min_k (Se_k - Sp_k)$$

В основу концепции построения индикатора вероятности положен следующий методический подход.

Прежде всего, индикатор должен отвечать требованиям:

- подавать достоверный сигнал продолжения (наличия) тренда (восходящий / нисходящий), что в свою очередь служит сигналом открытия позиции в соответствующем направлении;

- подавать предупреждающий сигнал смены текущей тенденции, что делает возможным интерпретацию сигнала - закрытие открытой позиции и открытия разворотной позиции в торговой стратегии.

Структура индикатора представляет собой общий индекс, описывающий множественную классификацию сигналов от различных известных индикаторов, осцилляторов и индексов различных показателей динамики валютного курса (переменных модели), выражающий совпадение сигналов для идентификации текущего динамического поведения цены инструмента.

Предположения модели:

- ни одна из переменных модели, ни их комбинация не зависят от переменной, которая моделируется (вероятность состояния рынка - типа текущего тренда);

- остатки в модели распределены логистически, функция распределения вероятности для линейной комбинации переменных и коэффициентов модели соответствует логистическому распределению [7].

Случайность и независимость объектов классификации достигается способом формирования выборки данных для построения модели.

Задача прогнозирования валютного курса рас-

падает на ряд подзадач, решение которых необходимо для достижения поставленной цели исследования.

1. Эмпирическая классификация динамического поведения валютного курса на интервале формирования выборки для построения эконометрической модели.

2. Качественный анализ факторов. Отобрать факторы, влияние которых на развитие тенденции валютного курса существенно.

3. Построение нескольких прогнозных моделей индикатора вероятности. Оценка параметров моделей.

4. Верификация моделей на тестовой выборке. Выбор лучшей в статистическом и экономическом смысле.

Решение указанных подзадач приведет к частному решению по достижению цели исследования ввиду множества вариантов построения моделей на 3-м этапе. Решение этой частной задачи рассматривается авторами в данной работе.

В качестве информационной базы используется история динамики торгов на спот – рынке EUR/USD, котировки рынка FoRex - часовые изменения курса EUR/USD с 2004 по 2009 годы. Для выгрузки исторических данных был использован раздел «Data bank» на сайте ДЦ «Альпари» (<http://www.alpari-idc.ru/ru/dc/databank.php>).

Для решения задач исследования использованы пакеты программ Statistica 6.0 и MS Excel 2007.

Исходная выборка формируется следующим образом: это таблица, в которой строки – ситуации соответствуют участкам тренда – восходящий / нисходящий, а столбцы – набор переменных (вектор X и их классификаций y).

Результаты исследования

Предварительная эмпирическая классификация сигналов с целью идентификации участков однодневного поведения цены закрытия инструмента и точек разворота текущей тенденции проводилась в три этапа.

1 этап. Анализ данных на основе критериев:

- последовательность более высоких максимумов и минимумов $max_{t+1} > max_t$,

$min_{t+1} > min_t$ (для $\forall t$), графический анализ - восходящий тренд;

- последовательность более низких максимумов и минимумов $max_{t+1} < max_t$,

$min_{t+1} < min_t$ (для $\forall t$), графический анализ - нисходящий тренд;

- неравенства $u_i < u_{i+1} > u_{i+2}$ или $u_i > u_{i+1} < u_{i+2}$ при выполнении одного из первых двух условий идентифицируют (соответственно) точку разворота промежуточной тенденции.

2 этап. Полученная классификация подтверждалась графическим анализом поведения следующих индикаторов и осцилляторов в *MetaTrader* (рис. 1):

- на часовом графике экспоненциальные средние разных порядков $EMA(5)$, $EMA(13)$, $EMA(21)$ определяют направление тренда, их пересечение – разворот. Если быстрая $EMA()$ выше медленной, то повышательная тенденция набирает скорость, если ниже медленной - понижательная тенденция набирает скорость;

- индикатор $MACD$ следует за трендом, пересечение оперативной и сигнальной линий – сигнал о развороте тенденции, поиск расхождений поведения цены и индикатора в качестве подтверждения сигнала;

- осциллятор RSI – пересечение уровней перекупленности / перепроданности - сигнал к развороту, поиск расхождений. Считается, что рынок, где осциллятор находится в зонах перекупленности / перепроданности, ждет консолидация или разворот [5];

- осциллятор $Stochastic$ – пересечение медленной и быстрой линий, пересечение уровней перекупленности / перепроданности - сигнал к разво-

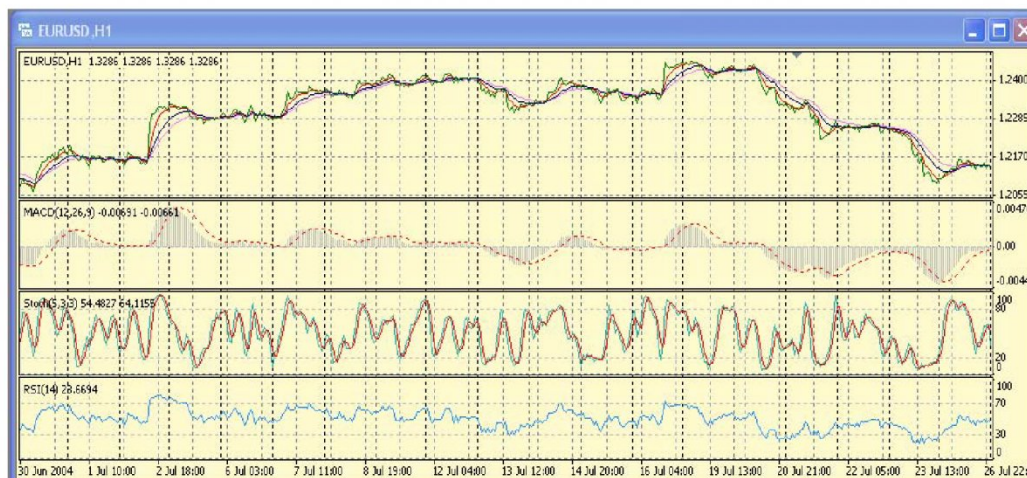


Рис. 1. Графический анализ поведения цены закрытия, скользящих средних разного периода, индикатора MACD и осцилляторов RSI и Stochastic

Таблица 1. Эмпирическая классификация

Начало интервала	Конец интервала	Классификация	Тип тренда
-	17.06.04 1:00	0	↓
17.06.04 2:00	17.06.04 23:00	1	↑
18.06.04 0:00	18.06.04 8:00	0	↓
18.06.04 9:00	21.06.04 1:00	1	↑
22.06.04 22:00	23.06.04 9:00	1	↑
23.06.04 10:00	23.06.04 19:00	0	↓
24.06.04 2:00	24.06.04 7:00	0	↓
24.06.04 8:00	24.06.04 21:00	1	↑
29.06.04 14:00	30.06.04 0:00	0	↓
30.06.04 9:00	30.06.04 23:00	1	↑
02.07.04 14:00	05.07.04 1:00	1	↑
07.07.04 6:00	07.07.04 14:00	1	↑
08.07.04 16:00	08.07.04 21:00	1	↑
13.07.04 2:00	13.07.04 8:00	0	↓
13.07.04 14:00	13.07.04 21:00	0	↓
14.07.04 0:00	14.07.04 19:00	1	↑
16.07.04 14:00	19.07.04 1:00	1	↑
20.07.04 12:00	21.07.04 5:00	0	↓
21.07.04 12:00	21.07.04 20:00	0	↓
23.07.04 11:00	26.07.04 2:00	0	↓
26.07.04 3:00	26.07.04 14:00	1	↑
27.07.04 15:00	28.07.04 1:00	0	↓
28.07.04 10:00	28.07.04 16:00	0	↓
03.08.04 15:00	03.08.04 18:00	1	↑
06.08.04 14:00	06.08.04 16:00	1	↑

Тип тренда - ↑ восходящий, ↓ нисходящий

роту, поиск расхождений в качестве подтверждения сигнала.

3 этап. Аналитическое подтверждение идентификации участков однородного поведения цены с помощью непараметрических критериев случайности: поворотных точек (9, с.486), инверсий (9, с. 491) и параметрического критерия *Аббе-Линника* (10, с. 517).

Если два индикатора (осциллятора) дают одинаковый сигнал и он подтверждается аналитиче-

скими критериями, принимается решение об отношении к определенному типу поведения цены закрытия торгового инструмента (восходящий, нисходящий, боковой тренд).

Эмпирическая классификация динамического поведения валютного курса на интервале формирования выборки для построения моделей прогноза представлена в табл.1 (112 наблюдений – нисходящий тренд, 152 наблюдения - восходящий тренд).

Качественный анализ факторов

Цель анализа - отобрать факторы, влияние которых на развитие тенденции валютного курса существенно.

На рынке FoRex доступна информация: цена открытия, цена закрытия, минимальная и максимальная цена за период, тиковый объем (характеризует активность рынка). Используем арифметические выражения показателей динамики цены закрытия, индикаторы и осцилляторы технического анализа, позволяющие решить главную задачу ТА: прогноз направления движения (разворот) рынка.

Практически все методы многомерного анализа в той или иной степени опираются на предположение о нормальном характере распределения входящих в модели случайных величин. В противном случае возникает необходимость подобрать преобразование переменных, приближающее исходный ряд к нормальному.

Известно, что любое непрерывное распределение монотонным непрерывным преобразованием случайной величины может быть приближено к нормальному, например, преобразования вида: $y = lg(x + a), a - const$ или

$$y = \ln\left(\frac{x - a}{x - b}\right), x \in (a, b)$$

которые также стабилизируют дисперсию [11]. В частности, преобразование Фишера [6] (при $n > 50$)

$$y = \frac{1}{2} \ln\left(\frac{1 + x}{1 - x}\right)$$

для $x \in (-1, 1)$ нормализует исходный ряд X . Пре-



Рис. 2. Матрица корреляций

Таблица 2. Оценка адекватности моделей

Модель	Обозначение	$\chi^2(v) / p$	Max likelihood	Уровень значимости χ^2 для коэффициентов
$L(x) = 6,0 + 11,1992 \cdot x_2 - 20,0 \cdot x_3 - 19,8217 \cdot x_1$	M1	$\chi^2(3) = 324,9306 / p=0,00000$	17,4835	P=0,00000; 0,05
$L(x) = 18,0056 \cdot x_4 - 7,6658 \cdot x_5 + 6,0 \cdot x_6 - 4,42162$	M2.2	$\chi^2(3) = 312,1601 / p=0,00000$	23,8688	P=0,00000; 0,01
$L(x) = 19,0 \cdot x_4 - 12,1016 \cdot x_5 - 1,0$	M2.1	$\chi^2(2) = 295,49 / p=0,0000$	32,2049	P=0,00000

образование вида $y = \sqrt[n]{x_i}$ стабилизирует дисперсию [6].

В результате предварительного анализа отобраны следующие факторы.

В качестве динамического показателя колеблемости относительно среднего уровня в момент времени t (амплитуда интенсивности валютных флуктуаций, волатильность доходности) принимаем линейное отклонение логарифмических приращений курсов валют $\sigma_r = |r_t - r_{cp}|$, где r_{cp} - среднее значение доходности за период n (последние $n=10$ периодов) [12]. Доходность определяется как логарифм индекса (коэффициента) роста цен $r_t = \ln\left(\frac{x_t}{x_{t-1}}\right)$, где x_t и x_{t-1} - цена закрытия текущего и предыдущего периодов.

В модели - это переменная *корЛинОткл*, к которой применено преобразование $y = \sqrt[n]{x_i}$ (для укрупнения исходного значения). Для восходящего / нисходящего тренда волатильность увеличивается с ростом / падением цены и уменьшается при корректировке цены, является подтвер-

ждающим индикатором в трендовых и разворотных моделях.

Индикатор *MACD* - схождение - расхождение скользящих средних, подтверждающий и опережающий индикатор, хорошо работает на трендах. Если $MACD > 0$ - восходящий, если $MACD < 0$ - нисходящий тренд. Формулы вычисления индикатора приведены в [5]. В качестве показателя сигнала точки разворота используем величину разности $d = \text{оперативная} - \text{сигнальная линия}$ (разница значений оперативной и сигнальной линий индикатора), $d \rightarrow 0$ - их пересечение дает точку разворота.

В модели - это переменная *КорMACDразн*, к переменной d применено преобразование $y = \sqrt[n]{x_i}$. Знак значения индикатора идентифицирует направление тренда: «+» - восходящий, «-» - нисходящий тренд, смена знака - точка разворота.

Осцилляторы *RSI* и *Stochastic* без запаздывания (а иногда и с опережением) показывает назревающие развороты трендов.

Рассмотрим индекс относительной силы *RSI*. Величина $dR = (RSI(8) - MA(8)) \rightarrow 0$, пересечение осциллятора со своей средней *MA(8)* предупреж-

Paramet.	Const.B0	КорMACDразн	RSIразПп	%Dnp
Const.B0	1,000000	-0,584904	-0,068924	-0,902252
КорMACDразн	-0,584904	1,000000	-0,477470	0,367814
RSIразПп	-0,068924	-0,477470	1,000000	0,212710
%Dnp	-0,902252	0,367814	0,212710	1,000000

Рис. 3. Матрица коэффициентов корреляции параметров модели M2.2

Таблица 3. Оценка моделей

Модель	R ²	Псевдо-R ²	Критерий Акайке	Критерий Шварца	Классификация
M1	0,922	0,06408	-1,1546	-1,11398	0 - 99%, 1 - 98%
M2.2	0,899	0,06259	-0,8418	-0,80117	0 - 95%, 1 - 98%
M2.1	0,864	0,06031	-0,6307	-0,60362	0 - 94%, 1 - 96%

дает о развороте цены. Формулы вычисления осциллятора приведены в [5].

В модели – это переменные $RSI_{разПП}$ и RSI_{100} . RSI_{100} – это величина $dR/100$, а $RSI_{разПП}$ – значение, полученное в результате применения преобразования Фишера к RSI_{100} . Знак значения индикатора идентифицирует направление тренда: «+» - восходящий, «-» - нисходящий тренд; смена знака и значение, близкое к нулю предупреждает о развороте тенденции.

Стохастический осциллятор *Stochastic* предсказывает разворот с большой точностью, состоит из 2 линий: медленной и быстрой. Формулы вычисления осциллятора приведены в [5].

Пересечение медленной и быстрой линий $d2=\%K - \%D \rightarrow 0$ предсказывает точку разворота [5, с. 109]. В модели – это переменная $\%Dnp$, значение которой получено в результате применения преобразования Фишера к переменной $\%D/100$.

Для идентификации точки разворота тенденции и областей тренда используем средний темп роста k_t^{sr} и прироста k_p цены закрытия за $n=10$ последних баров [12]:

$$k_t = \frac{P_t}{P_{t-1}}, k_t^{sr} = \sqrt[n]{k_{t1} \times \dots \times k_{tn}} \times 100\%,$$

$$k_p = k_t^{sr} - 100\%.$$

В модели – это переменная $корСрПрирост$, к которой применено преобразование $y = \sqrt[5]{x_i}$. Знак значения индикатора идентифицирует направление тренда: «+» - восходящий, «-» - нисходящий тренд; смена знака предупреждает о развороте тенденции. Для восходящего / нисходящего тренда средний прирост (цен закрытия за последние 10 периодов) увеличивается с ростом / падением цены и уменьшается при корректировке цены.

Для решения проблемы мультиколлинеарности переменных модели вычислены парные корреляции Пирсона. Результаты представлены в виде корреляционной матрицы (рис.2), уровень значимости коэффициентов $p < 0,05$.

Есть мнение, если основная задача модели – прогноз будущих значений результативного признака, то при $R^2 \geq 0,9$ (R^2 – коэффициент детерминации) наличие мультиколлинеарности не сказывается на прогнозных качествах модели [13].

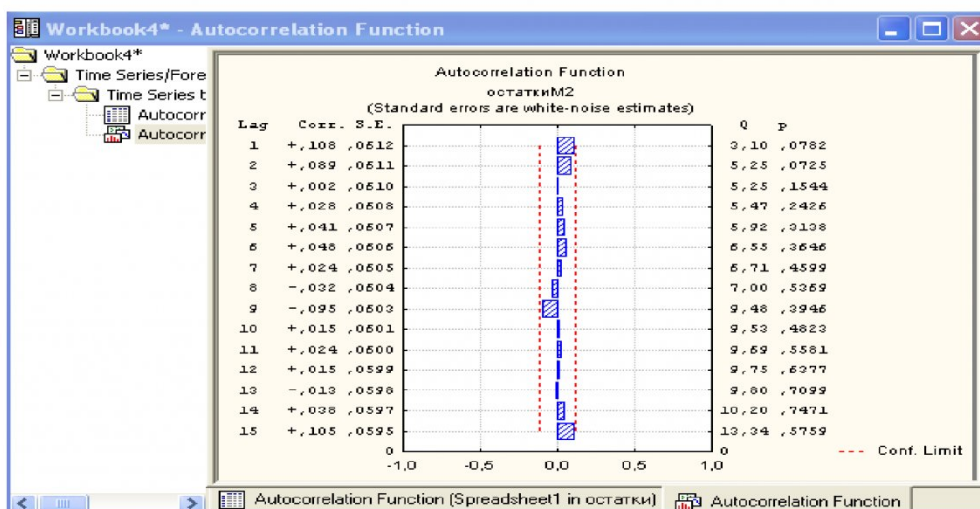


Рис. 4. Автокорреляционная функция остатков модели M2.2

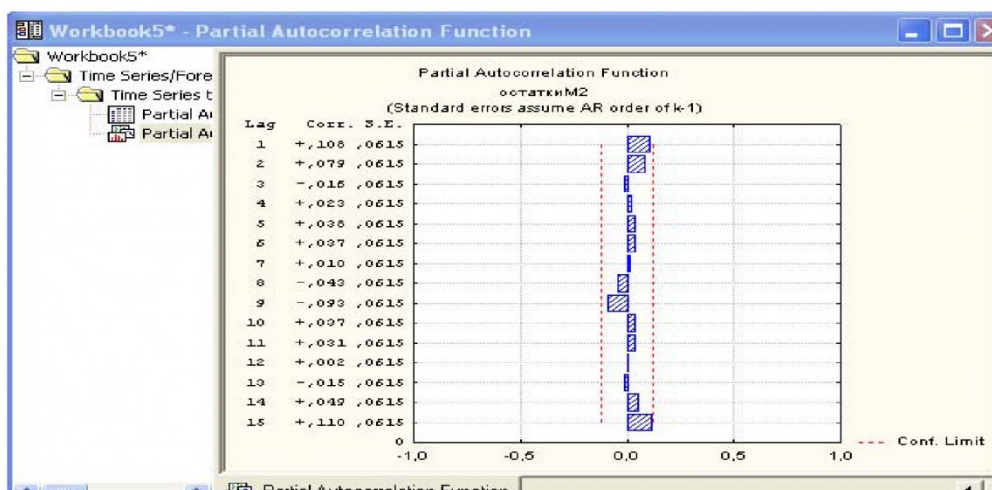


Рис. 5. Частная автокорреляционная функция остатков модели M2.2

Из таблицы видно, что переменные *КорМАСДразн* и *корСрПрирост* имеют средней степени значимую линейную связь ($\rho=0,73$). Остальные переменные имеют слабые ($\rho<0,25$) или умеренные ($\rho<0,5$) значимые линейные связи. Переменная *корЛинОткл* не имеет значимых линейных связей с другими переменными.

Построение моделей индикатора вероятности. Оценка параметров моделей.

Для доказательства адекватности моделей данным использованы следующие статистические критерии и процедуры [6, 7]:

- коэффициенты детерминации и псевдо - детерминации;
- *t*- критерий для оценки значимости коэффициентов и общей значимости модели в целом;
- проверка случайности остатков по критерию пиков;
- проверка отсутствия существенной автокорреляции 1-го порядка остатков модели по критерию Дарбина – Уотсона;
- критерии *Акайке* (AC) и *Шварца* (SW);
- графический анализ автокорреляционной и частной автокорреляционной функций остатков моделей.

Обозначим переменные: *корЛинОткл* – x_1 , *корСрПрирост* – x_2 , и *RSI_100* – x_3 , *КорМАСДразн* – x_4 , *RSIразПР* – x_5 , *%Dnp* – x_6 . Результаты подбора моделей с помощью программы *Statistica 6.0* представлены в табл. 2.

Модели оценены методом максимального правдоподобия, что предполагает нахождение оптимальных параметров с позиции максимума вероятности «попадания» модели в наблюдаемые значения. Уровень значимости p для соответствующей *t*-статистики для коэффициентов модели в большинстве случаев принимается равным $p=0,00000$. Для выявления общей значимости всех коэффициентов модели используется тест *log-likelihood*, когда логарифм функции правдоподобия от всех переменных сравнивается с логарифмом функции правдоподобия только от константы (табл. 2).

Структура модели *M2.1* отличается от *M2.2* отсутствием переменной x_6 , так как наблюдается сильная корреляция между переменной x_6 и свободным членом в модели *M2.2* (рис. 3), хотя «хорошие» прогнозные качества модели сохраняются (табл. 3). Высокий уровень значимости говорит об адекватности моделей данным.

Коэффициенты детерминации и псевдо - детерминации, значения критериев *Акайке* (AC) и *Шварца* (SW), результаты классификации моделей приведены в табл. 3.

По результатам анализа табл. 3 можно заключить: лучшей моделью является *M1*. У неё самый высокий коэффициент детерминации, самые низкие значения критериев *Акайке* и *Шварца*, а также высокий процент правильной классификации наблюдений. Модели *M2.2* и *M2.1* также имеют не-

плохие оценки. Классификация устойчива для всех моделей, порядок выбора объектов в программе *Statistica 6.0* не влияет на результат.

Проверка отсутствия существенной автокорреляции 1-го порядка остатков моделей осуществлялась по критерию Дарбина – Уотсона.

Для каждой модели вычислялась статистика Дарбина–Уотсона (*DW*) [14]. Верхние и нижние критические значения d_l и d_u двустороннего критерия при уровне значимости 0,05 для проверки нулевой гипотезы (отсутствие автокорреляции) взяты в [14, с. 223]. Получены результаты (число наблюдений $n>100$): для модели *M1* - $DW=1,4227$, число переменных $k=3$, $d_l = 1,55+...$ и $d_u=1,67+...$; для *M2.2* – $DW=1,54778$, число переменных $k=3$, $d_l = 1,55+...$ и $d_u=1,67+...$; для *M2.1* – $DW=1,25838$, число переменных $k=2$, $d_l = 1,57+...$ и $d_u=1,62+...$. Для всех моделей вычисленное значение статистики попадает в интервал $0 \leq DW < d_l$, что говорит о наличии положительной автокорреляции. Значения *DW* близки к границе d_l – области неопределенности.

Можно предположить, что автокорреляция вызвана усреднением данных по интервалам. Однако, проверка случайности остатков модели *M1* по критерию пиков показала справедливость нулевой гипотезы – остатки случайны. Значение критерия, вычисленное по наблюдениям выборки объемом $n=264$, $t_{факт} = 2,29473$, а $t_{теор} (p=0,01; n-1 = 263) = 2,334$. Фактическое значение критерия меньше теоретического $t_{факт} < t_{теор}$ – нулевая гипотеза принимается. Теоретическое значение кри-

Таблица 5. Интервалы тестирования моделей

Начало интервала	Конец интервала	Классификация	Тип тренда
21.06.04 2:00	21.06.04 18:00	0	↓
22.06.04 6:00	22.06.04 21:00	0	↓
23.06.04 20:00	24.06.04 1:00	0	↓
28.06.04 11:00	28.06.04 17:00	1	↑
01.07.04 0:00	01.07.04 2:00	1	↑
01.07.04 3:00	01.07.04 11:00	0	↓
05.07.04 3:00	05.07.04 18:00	0	↓
07.07.04 15:00	07.07.04 19:00	1	↑
08.07.04 22:00	09.07.04 8:00	1	↑
09.07.04 21:00	12.07.04 1:00	1	↑
13.07.04 9:00	13.07.04 22:00	0	↓
14.07.04 22:00	15.07.04 11:00	0	↓
19.07.04 2:00	19.07.04 6:00	1	↑
21.07.04 8:00	21.07.04 11:00	1	↑
21.07.04 12:00	21.07.04 20:00	0	↓
23.07.04 3:00	23.07.04 10:00	0	↓
27.07.04 9:00	27.07.04 14:00	1	↑
30.07.04 4:00	30.07.04 15:00	1	↑
02.08.04 3:00	02.08.04 13:00	1	↑
02.08.04 14:00	03.08.04 10:00	0	↓
03.08.04 11:00	03.08.04 14:00	1	↑
03.08.04 23:00	04.08.04 14:00	0	↓
04.08.04 15:00	05.08.04 1:00	1	↑

Тип тренда - ↑ восходящий, ↓ нисходящий

терия взято из таблицы [15, с. 130]. Графический анализ автокорреляционной и частной автокорреляционной функций для остатков модели *M2.2* (рис. 4 и рис. 5) показывает отсутствие корреляции. Графический анализ автокорреляционной и частной автокорреляционной функций для остатков модели *M2.1* показывает наличие несущественной положительной автокорреляции.

ROC – анализ логит – моделей

Для построения ROC - кривой используется алгоритм, предложенный в [8].

Почти прямой верхний левый угол ROC – кривых для исследуемых моделей указывает на большую предсказательную способность моделей.

Пороги отсекаемых моделей, найденные по двум критериям - максимальной суммарной чувствительности и специфичности, баланса между чувствительностью и специфичностью, представлены в табл. 4. Для модели *M2.1* можно использовать два порога отсекаемых: $\zeta_1 \approx 0,22$ и $\zeta_2 \approx 0,65$.

Верификация моделей на тестовой выборке. Выбор лучшей модели в статистическом и экономическом смысле.

Проверка адекватности моделей осуществляется по тем свойствам, которые считаются существенными для исследования, и, прежде всего, проверка устойчивости работы моделей к случайным изменениям на рынке на историческом интервале данных вне выборки построения моделей (метод тестирования за пределами выборки).

Пригодными для прогнозирования валютного курса признаются модели статистически значимые, с оптимальным порогом отсекаемых, устойчивые к случайным изменениям на рынке. Используются критерии: % угаданных направлений и оценка прибыли – количество пунктов, на которое изменится валютный курс при верном прогнозе и убытка – при ошибочном прогнозе направления тренда.

Интервалы тестирования моделей (вне выборки построения моделей) представлены в табл. 5 (214 наблюдений).

Оценивается простейшая торговая стратегия:

- вход на покупку по сигналу – прогнозное значение вероятности $> \zeta_m$ (тренд восходящий);
- вход на продажу по сигналу - прогнозное значение вероятности $< \zeta_m$ (тренд нисходящий);
- смена соответствующего сигнала – закрытие текущей и открытие разворотной позиции, ζ_m – тестируемое значение порога отсекаемых соответствующей модели;
- стоп - лоссы и трейлинг – стопы не используются.

Тестировались следующие пороги: для *M1* - $\zeta_m = 0,5$; для *M2.2* - $\zeta_m = 0,75$; для *M2.1* два порога - $\zeta_m^1 = 0,66$ и $\zeta_m^2 = 0,22$.

Для пары *EUR/USD* прибыль (убыток) в пунктах при покупке равна разнице между ценой закрытия и ценой открытия.

Прибыль (убыток) в пунктах при продаже равна разнице между ценой открытия и ценой закрытия (*1 пункт = 0,0001*).

В табл. 6 приведены результаты торговли и процент правильных прогнозов для каждой модели. Наилучший результат - наибольшая прибыль получен для модели *M2.1*, несколько хуже – для модели *M2.2*. По критерию процент правильных прогнозов лучшей моделью оказалась *M2.2* и модель *M2.1* для порога отсекаемых $\zeta_m^2 = 0,22$.

В соответствии с сформулированными выше критериями, пригодными для прогнозирования валютного курса, устойчивые к случайным изменениям на рынке, признаются все три модели. Лучшие результаты (в том числе с позиции полученной прибыли) показала модель *M2.1* с порогом отсекаемых $\zeta_m^2 = 0,22$.

В табл. 7 представлены результаты теста на чувствительность переменных моделей. Исследо-

Таблица 6. Верификация моделей на интервале тестирования

Модель (N=214)	Обозначение	Верно	Ошибка	% верных прогнозов	Прибыль/ Убыток
$L(x) = 6,0 + 11,1992 \cdot x_2 - 20,0 \cdot x_3 - 19,8217 \cdot x_1$	M1	195	19	91,12%	414/-38 (+376)
$L(x) = 18,0056 \cdot x_4 - 7,6658 \cdot x_5 + 6,0 \cdot x_6 - 4,42162$	M2.2	207	7	96,73%	344/-64 (+280)
$L(x) = 19,0 \cdot x_4 - 12,1016 \cdot x_5 - 1,0$	M2.1 порог 1	203	11	94,86%	522/-47 (+475)
	порог 2	206	8	96,26%	

Таблица 7. Анализ чувствительности переменных моделей

Модель/ переменная	Модель 1	Модель 2.2	Модель 2.1	Среднее	Стандартное отклонение
X_2	0,994407415*	---	---	0,047415	0,525283515
X_3	-5,24521E-06	---	---	-0,01603	0,15572679
X_1	0,941229224*	---	---	0,236238	0,047305062
X_4	---	0,96454668*	0,49226028	0,059257	0,223011225
X_5	---	-8,4043E-06	0,49998361	-0,01643	0,159343138
X_6	---	0,84840292*	---	0,66343	0,416903794

валось влияние изменения каждой переменной соответствующей модели на два стандартных отклонения от среднего уровня (уменьшение и увеличение на одно стандартное отклонение). Существенное (наибольшее) влияние на результаты – вероятность роста (падения) тренда оказывают переменные, отмеченные знаком «*». Это линейное отклонение и средний прирост – показатели волатильности цены, а также разность значений оперативной и сигнальной линий индикатора *MACD*, знак величины которой, показывает направление (и смену) тренда.

Выводы

На основе нового подхода к прогнозированию

валютного курса получены пригодные для использования в торговле пары *EUR/USD* модели логит - регрессии, статистически значимые, с оптимальным порогом отсека, устойчивые к случайным изменениям на рынке.

Разработанные модели индикатора вероятности прогноза направления движения цены инструмента *EUR/USD* в дальнейшем будут использоваться для конструирования эффективных торговых стратегий на валютном рынке *Forex*, что позволит повысить обоснованность принятия решений об инвестировании средств в рынок российскому частному трейдеру и другим категориям инвесторов.

СПИСОК ЛИТЕРАТУРЫ

1. *Pring, Marlin J.* Technical Analysis Explained. 3-rd edition. - New York: McGraw-Hill, 1991.
2. *Эрлих, А.* Технический анализ товарных и финансовых рынков. Прикладное пособие.- М.: ИНФРА-М, 1996.
3. *Мерфи, Д.* Межрыночный технический анализ.- М.: Диаграмма, 1999.
4. *Сафонов, В.* Практическое использование волн Эллиотта в трейдинге: диагностика, прогнозирование и принятие решений.- М.: Альпина Паблишер, 2002.
5. *Якимкин, В.* Как начать зарабатывать на валютном рынке *Forex*. – М.: СмартБук, 2008.
6. *Ллойд, Э.* Справочник по прикладной статистике / Э. Ллойд, У. Ледерман.- М.: Финансы и статистика, 1990.- Т. 1.
7. *Драйпер, Н.* Прикладной регрессионный анализ / Н. Драйпер, Г. Смит. - М.: Финансы и статистика, 1986.- Кн. 1.
8. *Zweig, M.H.* ROC Plots: A Fundamental Evaluation Tool in Clinical Medicine / *Zweig M.H., Campbell G.* // *Clinical Chemistry*, 1993.-V. 39, No. 4.
9. *Кендалл, М. Дж.* Многомерный статистический анализ и временные ряды/ М. Дж. Кендалл, А. Стьюарт. - М.: Наука, Главная редакция физ.-мат. литературы, 1976.
10. *Кобзарь, А.И.* Прикладная математическая статистика. Для инженеров и научных работников. - М.: Физматлит, 2006.
11. *Айвазян, С.А.* Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное издание / Айвазян С.А, Енюков И.С., Мешалкин Л.Д. – М.: Финансы и статистика, 1983.
12. Статистика: учеб. пособие/ Харченко Л.П., В.Г. Долженкова [и др.] – Изд. 2-е перераб. и доп. – М.:ИНФРА-М, 2005.
13. *Просветов, Г.И.* Статистика: Задачи и решения: учеб.-прак. пособ.- М.: Альфа-пресс, 2008.
14. *Кейн, Э.* Экономическая статистика и эконометрия. Введение в количественный экономический анализ / Вып. 2.- М.: Статистика, 1977.
15. *Закс, Л.* Статистическое оценивание. - М.: Статистика, 1976.

□ Авторы статьи:

Крюкова
Валентина Валентиновна
– канд. техн. наук, доц. каф. вычислительной техники и информационных технологий КузГТУ,
Email: kvv.vt@kuzstu.ru.

Крюков
Павел Алексеевич
– зам. управляющего Филиала «Кемеровский» ОАО «Собинбанк»,
Email: kpa.2008@mail.ru.