УДК 330.43

Нечеткие параметры финансовых временных рядов

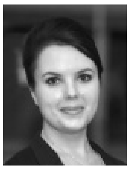
Аннотация. *Авторы анализируют возможность применения методов теории нечетких множеств для вычисления коэффициентов альфа и бета и для прогнозирования стоимости акций. С помощью нечеткой линейной регрессии проводится оценивание коэффициентов альфа и бета для фонда Bull & Bear U.S. & Overseas по месячным данным за январь 1996 г. — декабрь 1998 г., и результаты сравниваются с оценками, полученными при использовании обычной линейной регрессии. Применение аппарата теории нечетких множеств для прогнозирования финансовых временных рядов позволяет частично автоматизировать процесс технического анализа в области распознавания фигур и определения тенденций.*

Ключевые слова: *нечеткая линейная регрессия; альфа-коэффициент; бета-коэффициент; нечеткие временные ряды; прогнозирование стоимости акции.*

Abstract. *The possibility of applying the methods of fuzzy set theory for calculating alfa and beta coefficients and forecasting time series is discussed. Fuzzy linear regression is applied to estimate alpha and beta factors for the Bull & Bear US & Overseas fund for the monthly data from January 1996 to December 1998. The results are compared with estimates obtained using conventional linear regression. Applying fuzzy sets to analysis of financial time-series allows to partly automatize the process of technical analysis in the field of pattern recognition and identification of trends.*

Keywords: *fuzzy linear regression; alpha coefficient; beta coefficient; fuzzy time series; stock price forecasting.*

Веденеев Д.А.,
студент магистратуры
Финансового университета
✉ ridatri@yandex.ru



Петрова М.В.,
студентка магистратуры
Финансового университета
✉ m.vl.petrova@bk.ru

Далеко не всегда неопределенность, связанная с экономическими явлениями, имеет вероятностный характер. Применение классических вероятностных методов может привести к ошибкам, недооценке рисков или неадекватной интерпретации полученных количественных оценок. Теория нечетких множеств основывается на более слабых предположениях относительно характера неопределенности и зачастую оказывается подходящим инструментом моделирования экономических явлений (см., например, [1]). Применительно к теории нечетких множеств мы следуем терминологии, принятой в книге [2]. Расхождения оговариваются особо.

1. Оценивание коэффициентов α и β

1.1. Четкая линейная регрессия. В рамках модели CAPM справедливо следующее соотношение для доходности R:

$$E(R - R^f) = \beta E(R^M - R^f), \qquad (1)$$

где R^f — ставка безрискового актива, R^M — доходность рыночного портфеля,

Научный руководитель: ***Гисин В.Б.,*** *кандидат физико-математических наук, профессор, заведующий кафедрой «Математика».*

$$\beta = \frac{Cov(R^M, R)}{V(R^M)} = \frac{Cov(R^M - R^f, R - R^f)}{V(R^M - R^f)}.$$

В качестве аппроксимации рыночного портфеля при тестировании обычно используют некоторый индекс, включающий довольно большое число акций.

«Нормальной» доходностью является величина $E(R) = R^f + \beta E\,(R^M - R^f)$.

Значения, превышающие эту величину, считаются «сверхнормальными» доходностями.

Коэффициент β в (1) совпадает с коэффициентом β в регрессии:

$$R - R^f = \alpha + \beta(R^M - R^f) + \varepsilon. \qquad (2)$$

В классических моделях предполагается, что ε — нормально распределенная случайная величина, а ошибки для отдельных наблюдений некоррелированные. На российском фондовом рынке (да и на более развитых фондовых рынках) эти условия могут нарушаться, и применение классических методов в такой ситуации выглядит недостаточно обоснованным [3, 4, 5, 6]. Более того, классические подходы могут привести к недооценке рисков [4, 6].

Коэффициент бета оценивает чувствительность цен акций к фондовому индексу, т.е. показывает, как изменится цена акции при изменении значения индекса. Значение коэффициента бета равное нулю означает, что изменчивость цены акции практически не зависит от изменчивости индекса. Значение коэффициента больше нуля говорит о положительной корреляции между ценой данной акции и индекса, т.е. рост индекса (или его падение) сопровождается ростом (или падением) цены акции. Отрицательная бета означает, что цена акции меняется в направлении, противоположном изменению индекса.

Оценку $\hat{a}$ можно получить, оценивая коэффициент регрессии (2) по предыстории. Эта оценка называется α-коэффициентом Йенсена и служит для оценивания успеха данного актива на рынке. Если $\hat{a} > 0$, то актив показывает лучшую доходность по сравнению с «нормальной» (и, следовательно, его стоит включить в свой портфель). Значение $\hat{a} \approx 0$ соответствует «нормальной» доходности, а $\hat{a} < 0$ соответствует доходности меньшей, чем нормальная.

Рассмотрим пример из книги [7, с. 465–470]. Часто менеджеры взаимных фондов утверждают, что, благодаря их знаниям, опыту и искусству управления портфелем, фонд (и его вкладчики) получает доход выше, чем среднерыночный. Данное утверждение можно тестировать в рамках описанной выше однофакторной модели (2), оценивая α-коэффициент Йенсена.

Ниже приведены четкие оценки такого рода для фонда *Bull & Bear U.S. & Overseas*. В качестве безрискового актива взяты 5-летние государственные облигации США, а в качестве рыночного портфеля — *MSCI* индекс США. В *таблице 1* приведен результат оценивания модели (1) с помощью метода наименьших квадратов (функция ЛИНЕЙН, *MS Excel*), оценивание производилось на интервале с января 1996 г. по декабрь 1998 г. для месячных доходностей.

Таблица 1

Стандартная МНК-оценка параметров регрессии

BBUO — R^f	USA — R^f
β = 1,140616	α = –1,25443
σ_β = 0,119694	σ_α = 0,708935
R^2 = 0,727586	3,405385
90,81016	34
1053,094	394,2861

Таким образом, регрессия (2) принимает вид:

$$R - R^f = -1{,}25 + 1{,}14(R^M - R^f).$$

Видим, что α-коэффициент Йенсена отрицателен и значимо отличается от нуля. Следовательно, на этих данных и рассмотренном периоде времени утверждение менеджеров не подтверждается реальными данными: доходность фонда меньше, чем «нормальная».

Коэффициент бета положителен и равен 1,14, значит, при росте (соответственно падении) цены индекса на 1% цена актива вырастет (соответственно упадет) на 1,14%.

1.2. Нечеткая линейная регрессия. Рассмотрим теперь случай нечеткой регрессии (2):

$$R - R^f = A + B(R^M - R^f), \qquad (3)$$

где $A = (a - p,\ a,\ a+p)$, $B = (b - q,\ b,\ b+q)$ — симметричные треугольные нечеткие числа, выполняющие роль нечетких аналогов коэффициентов альфа и бета соответственно [8].

Для упрощения обозначений положим $R - R^f = Y$, $R^M - R^f = x$, тогда (3) примет вид:

$$Y_j = A + Bx_j, \qquad (4)$$

где x_j — четкие числа, полученные в результате наблюдений, а Y_j — симметричное треугольное нечеткое число. Несложное вычисление показывает, что

$$Y_j = (a_j + bx_j - (p + q|x_j|),\ a_j + bx_j,\ a_j + bx_j + (p + q|x_j|)).$$

Найдем теперь суммарную нечеткость в определении выходных значений при m имеющихся наблюдениях:

$$r = mp + \sum_{j=1}^{m} q\left|x_j\right| = mp + q\sum_{j=1}^{m}\left|x_j\right|.$$

Задача нечеткой линейной регрессии сведется к следующей задаче линейного программирования:

$$r = mp + q\sum_{j=1}^{m}\left|x_j\right| \to min; \qquad (5)$$

$$y_j \geq (a + bx_j) - (1 + h)(p + q|x_j|), j = 1, 2, ..., m;$$

$$y_j \leq (a + bx_j) - (1 + h)(p + q|x_j|), j = 1, 2, ..., m;$$

$$p \geq 0, q \geq 0.$$

Решаем поставленную задачу оптимизации для данных, описанных в предыдущем пункте, с помощью «Поиска решения» в MS Excel. Для уровня h = 0 решение получается при a = –0,79, b = 1,09, p = 6,45, q = 0, т.е.

$$Y = (-7{,}74;\ -0{,}79;\ 5{,}66) + 1{,}09x. \qquad (6)$$

Видим, что β-коэффициент является четким числом, равным 1,09, что не сильно отличается от его оценки, полученной с помощью четкой регрессии (β = 1,14). Вся неопределенность связана с α-коэффициентом Йенсена, который имеет довольно большой интервал толерантности, но наиболее возможное значение –0,79 согласуется с предыдущими результатами (отлично от нуля и отрицательно).

Применение нечеткой линейной регрессии позволяет дать более адекватную интерпретацию количественным оценкам, оставляя неопределенность там, где она объективно имеется.

2. Прогнозирование стоимости акций на основе нечетких временных рядов

Порой кажется, что скачки цен на бирже представляют собой некое хаотичное движение, которое абсолютно невозможно предсказать. Однако существует большое количество методов, которые позволяют «упорядочить» и объяснить динамику цен, в том числе и неклассических [9]. Вполне естественно, что у каждого из этих методов есть свои плюсы и минусы. Мы рассмотрим способ построения прогноза, основанный на анализе нечетких временных рядов [10, 11, с. 553–564].

2.1. Исходные данные. В качестве объекта исследования используем акции американской компании *CA Technologies*, которая занимается разработкой программного обеспечения. Акции компании торгуются на бирже *NASDAQ*.

Рассмотрим в качестве исходных данных стоимость акций со 2 февраля 2015 г. по 14 апреля 2015 г.

Стоит отметить, что прогнозироваться будет не сама стоимость акции в будущий период времени, а изменение стоимости, поэтому из исходных данных получим новый временной ряд, состоящий из 49 значений.

2.2. Алгоритм прогнозирования. Процесс получения прогнозных значений можно разбить на 8 шагов.

1. Определяется универсальное множество U, которое покрывает все исторические данные, т.е. интервал между наименьшим и наибольшим значением в рассматриваемой выборке. Обычно выбирают такое множество, которое может быть чуть больше минимального, но которое позволяет производить расчеты в более удобной форме.

Для рассматриваемых данных удобнее всего выбрать интервал [–1,045; 0,845] в качестве универсального множества.

2. Универсальное множество делится на некоторое количество равных интервалов, каждый из которых соответствует определенному лингвистическому терму A_i.

Разобьем множество значений на 9 равных интервалов:

$A_1 = [-1{,}045; -0{,}835]$ — очень большое падение;
$A_2 = [-0{,}835; -0{,}625]$ — большое падение;
$A_3 = [-0{,}625; -0{,}415]$ — падение;
$A_4 = [-0{,}415; -0{,}205]$ — слабое падение;
$A_5 = [-0{,}205; 0{,}005]$ — слабые изменения;
$A_6 = [-0{,}005; 0{,}215]$ — слабый рост;
$A_7 = [0{,}215; 0{,}425]$ — рост;
$A_8 = [0{,}425; 0{,}635]$ — большой рост;
$A_9 = [0{,}635; 0{,}845]$ — очень большой рост.

3. Формируется последовательность отношений следования «на день вперед» $A_i \to A_j$.

Для построения последовательности отношений попарно рассматриваются последовательные фаззифицированные данные.

Например, значение изменения стоимости акции 10 марта попадает в интервал A_3, а на следующий день, 11 марта, значение принадлежит интервалу A_5. Таким образом, формируется отношение $A_3 \to A_5$.

4. Отношения, полученные на предыдущем шаге, образуют группы, в которые входят все отношения с одинаковыми левыми частями. Отношения $A_1 \to A_5$ и $A_1 \to A_6$ объединяются в группу $A_1 \to A_5, A_6$. Говоря об экономическом смысле данных групп, можно сказать, что они отражают те вероятные изменения стоимости акций, которые могут следовать после определенного изменения.

Так, для группы $A_1 \to A_5, A_6$ можно сделать вывод, что после очень большого падения стоимости имеют место слабые изменения или слабый рост.

Из получившихся групп формируются отношения R_i по правилу:

$$R_i = A_i^T \times A_j,$$

где j — номера тех термов, которые следуют за термом A_i.

Для рассмотренной ранее группы $A_1 \to A_5, A_6$ получается следующее отношение R_1:

$$R_1 = A_1^T \times A_5 \cup A_1^T \times A_6.$$

На основе полученных отношений формируется матрица R согласно следующим правилам:

- элемент r_{ij} равен 1, если в отношении R_i присутствует член $A_i^T \times A_j$;
- элемент r_{ij} равен 0,5; если член $A_i^T \times A_j$ не присутствует, однако есть член(ы) $A_i^T \times A_{j\pm 1}$;
- элемент r_{ij} равен 0, если ничего из вышесказанного не выполняется.

Рассмотрим процесс построения матрицы R более подробно.

Допустим, что в рассматриваемых данных была пара последовательных изменений $A_3 \to A_4$, тогда элемент $r_{3,4}$ матрицы R будет равен 1. Предположим также, что пара $A_3 \to A_3$ в данных отсутствует, однако из-за наличия «соседней» пары $A_3 \to A_4$ элемент $r_{3,3}$ матрицы R будет равен 0,5. И наконец, если нет ни пары $A_3 \to A_2$, ни «соседних» пар $A_3 \to A_1$ или $A_3 \to A_3$, то элемент $r_{3,3}$ будет равен 0. В этом случае получится следующая матрица R:

$$R = \begin{pmatrix} 0 & 0 & 0 & 0{,}5 & 1 & 1 & 0{,}5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0{,}5 & 1 & 1 & 1 & 0{,}5 & 0 & 0 \\ 0 & 0 & 0{,}5 & 1 & 0{,}5 & 1 & 1 & 0{,}5 & 0 \\ 1 & 0{,}5 & 1 & 0{,}5 & 1 & 1 & 1 & 1 & 0{,}5 \\ 1 & 0{,}5 & 1 & 1 & 1 & 0{,}5 & 1 & 0{,}5 & 1 \\ 0 & 0{,}5 & 1 & 1 & 1 & 0{,}5 & 1 & 0{,}5 & 0 \\ 0 & 0 & 0 & 0{,}5 & 1 & 1 & 0{,}5 & 1 & 0{,}5 \\ 0 & 0 & 0 & 0{,}5 & 1 & 0{,}5 & 1 & 0{,}5 & 0 \end{pmatrix}.$$

5. Производится фаззификация известных данных, т.е. перевод четких чисел в нечеткие. Степень принадлежности μ некоторого известного значения x_t интервалу i определяется с помощью формулы:

$$\mu_i(x_t) = \frac{1}{1 + k(x_t - u_i)^2},$$

где x_t — фактическое значение временного ряда в момент времени t; u_i — середина интервала i; k — положительная постоянная, которая отражает, насколько строго относится исследователь к принадлежности значений интервалам (*рис. 1*).

Можно заметить, что при увеличении k график сжимается, т.е. степень принадлежности уменьшается.

Из получившихся значений степени принадлежности формируется матрица M, каждая строка которой отражает то, в какой степени соответствующее этой строке значение временного ряда принадлежит одному из термов A_i, которые в свою очередь расположены в столбцах:

$$M = \begin{pmatrix} \mu_1(x_1) & \dots & \mu_9(x_1) \\ \mu_1(x_2) & \cdots & \mu_9(x_2) \\ \vdots & \ddots & \vdots \\ \mu_1(x_n) & \cdots & \mu_9(x_n) \end{pmatrix}.$$

6. Прогнозное значение изменения V_{t+1} будем находить с помощью min-max композиции:

$$V_{t+1} = M_t \circ R.$$

7. Наконец, необходимо дефаззифицировать полученные прогнозные значения. Для этого применим следующую формулу:

$$\hat{x}_i = \sum_{j=1}^{9} v_{ij} \cdot u_j,$$

где u_j — середина j-го интервала.

8. Найдем прогнозные значения стоимости акций:

$$y_{t+1} = y_t + x_t,$$

где y_{t+1} — прогнозное значение стоимости акции в момент $t+1$; y_t — реальные данные в момент t; x_t — прогнозное значение изменения стоимости акции.

2.3. Результаты прогнозирования. В результате исследования был получен временной ряд прогнозных значений, который вместе с реальными данными изображен на *рис. 2*.

Максимальная ошибка прогноза составила приблизительно 3,75%. Возможно, одним из факторов, которые повлияли на расхождение прогнозных и реальных данных, является неоптимальный выбор постоянной k, которая применяется в формуле для расчета степени принадлежности.

В процессе исследования были выявлены некоторые плюсы и минусы применения методов теории нечетких множеств. К положительным сторонам можно отнести то, что используются логически понятные и естественные конструкции. Нечеткие модели имеют значительный ресурс гибкости при настройке и позволяют учитывать слабо формализованное мнение эксперта. Это позволяет надеяться, что методы те-

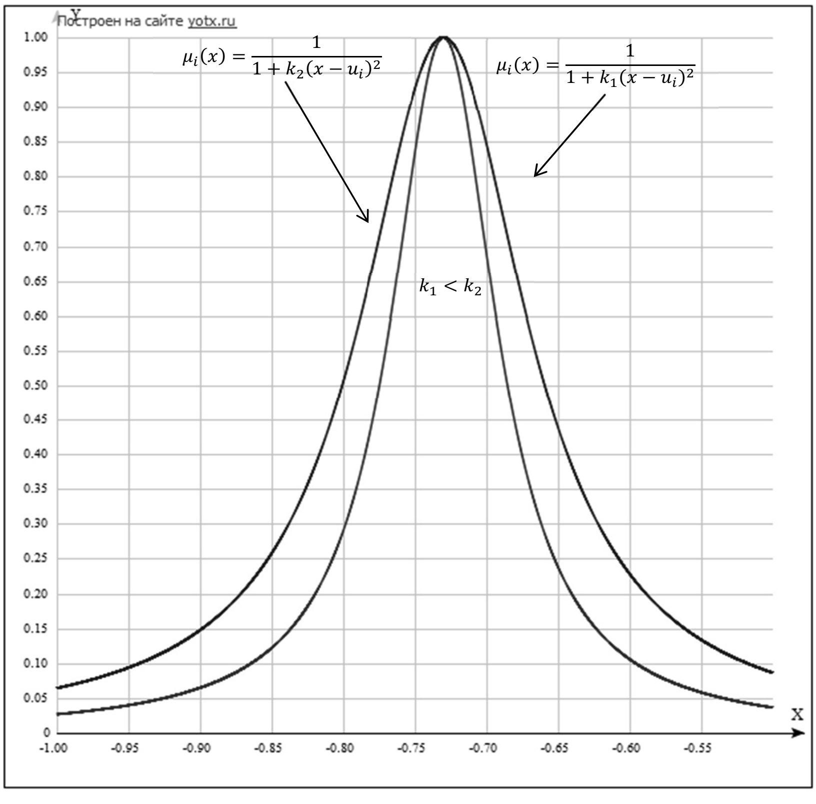


Рис. 1. **Функция принадлежности нечеткому интервалу**

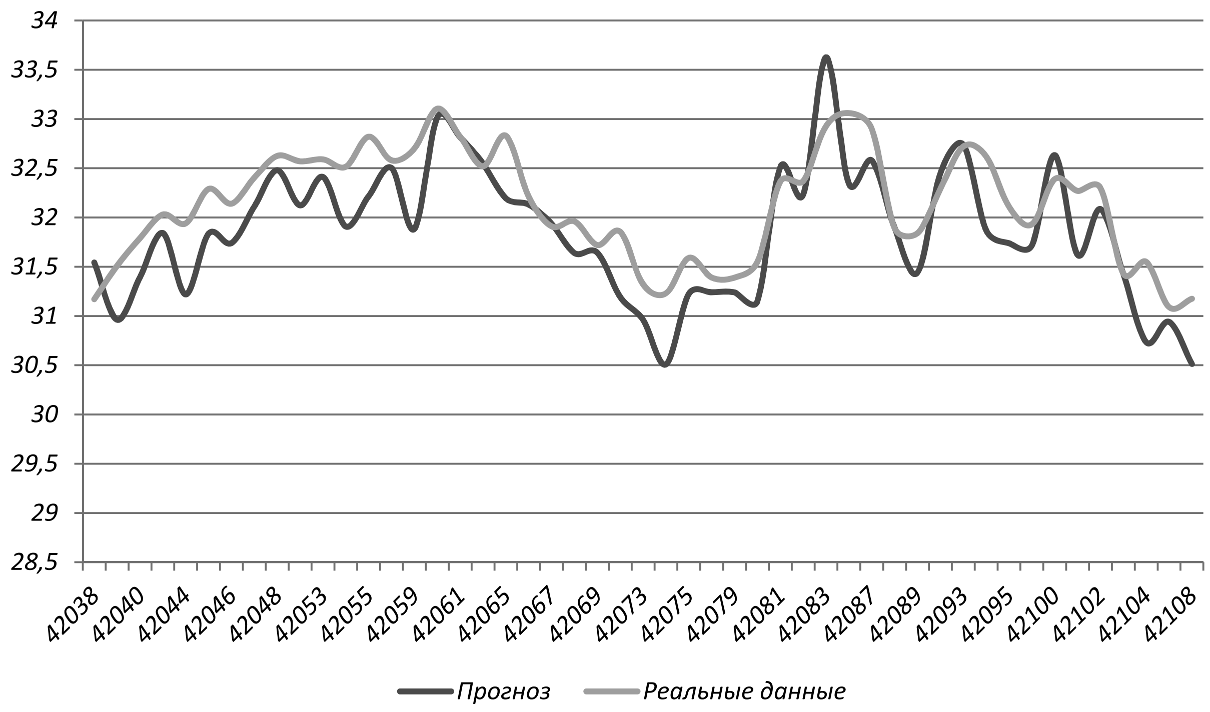


Рис. 2. **Прогноз и реальные данные**

ории нечетких множеств будут применяться и в других задачах со слабо формализованными данными [12].

К негативным сторонам можно отнести некоторую громоздкость и нетривиальность вычислений, особенно тех, которые связаны с выполнением неарифметических операций (например, вычисление минимаксной композиции матриц). Кроме того, неясно, как оценивать точность получаемых результатов. В то же время неточность результатов, вероятно, отражает неопределенность данных и дает в некотором смысле «честный» ответ на вопрос в ситуации неполной и неточной информации.

Литература

1. Волкова Е.С., Гисин В.Б. Меры возможности и внутренняя норма доходности инвестиционных проектов с нечетко определенными платежами // Вестник Финансового университета. 2014. № 3 (81). С. 93–104.
2. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб: БХВ-Петербург, 2005. 736 с.
3. Гисин В.Б., Коннов В.В., Шаров В.Ф. Вероятностные модели для анализа ценообразования активов на фондовых рынках: Монография. М.: Финансовый ун-т, 2012. 152 с.
4. Гисин В.Б., Марков А.А. Ценообразование производных инструментов европейского типа на фрактальном рынке с транзакционными издержками // Вестник Финансового университета. 2011. № 1 (61). С. 34–41.
5. Гисин В.Б. Оценка стоимости опционов на рынке с трансакционными издержками и негауссовой ценовой динамикой // Экономические науки. 2013. № 100. С. 165–168.
6. Гисин В.Б., Ярыгина И.З. Управление рисками стоимости активов в динамических моделях рынка с трансакционными издержками // Вестник Института экономики Российской академии наук. 2014. № 1. С. 85–101.
7. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. М.: Дело, 2004. 504 с.
8. Волкова Е.С., Гисин В.Б. Нечеткая линейная регрессия в оценке недвижимости // Вопросы оценки. 2015. № 1 (79). С. 26–33.
9. Шкляев А.О. Прогнозирование финансовых временных рядов методом скрытых марковских моделей // Научные записки молодых исследователей. 2015. № 1. С. 17–21.
10. Song Q., Chissom B.S. Forecasting enrollments with fuzzy time series — part 1 // Fuzzy Sets and Syst. 1993. No. 54. C. 1–9.
11. Şah M., Degtiarev K. A new trend heuristic time-variant fuzzy time series method for forecasting enrollments // Computer and Information Sciences-ISCIS 2005. Springer Berlin Heidelberg, 2005.
12. Михалькевич И.С. Повышение достоверности слабоформализованных данных // Научные записки молодых исследователей. 2014. № 2. С. 17–21.