

УДК 336.76(045)

ПОСТРОЕНИЕ ТОРГОВОЙ СТРАТЕГИИ НА ОСНОВЕ МЕТОДОВ НЕЧЕТКОЙ ЛОГИКИ

Ратников А.А.,

студент, Финансовый университет, Москва, Россия
aa_ratnikov@mail.ru

Аннотация. В статье используются нечеткая логика и методы теории нечетких множеств в построении торговой стратегии для акций компании Microsoft. Алгоритм основан на применении лингвистических переменных и нечеткого вывода Мамдани. Все вычисления выполняются в математическом пакете R. Актуальность темы обусловлена тем, что применение нечетких множеств позволяет описать математическим языком оценки финансовых показателей и индикаторов, которые использует трейдер на основе своего субъективного опыта. Таким образом, нечеткая логика естественно вписывается в теорию алготрейдинга, поскольку позволяет реализовать идеи трейдера, одновременно устраняя из процесса торговли человеческий фактор. В ряде математических пакетов создан и реализован весь необходимый для данной прикладной задачи математический аппарат (задание нечетких множеств, лингвистических переменных, нечеткий вывод, фаззификация и дефаззификация).

Ключевые слова: нечеткая логика; мягкие вычисления; лингвистические переменные; алгоритм Мамдани; фондовый рынок; торговая стратегия; алготрейдинг

CREATING A TRADING STRATEGY BASED ON FUZZY LOGIC

Ratnikov A.A.,

student, Financial University, Moscow, Russia
aa_ratnikov@mail.ru

Abstract. The article is devoted to the methods of fuzzy logic in the trading strategy for shares of Microsoft. The strategy is based on the linguistic variables and Mamdani's fuzzy inference method. The R language is used for all calculations.

Subject of this research is interesting because methods of fuzzy logic enable to describe trader's assessment of financial indicators in mathematical way. Thus, fuzzy logic so naturally fits into the theory of algorithmic trading in that it makes it possible to realize any ideas of trader as an unflappable algorithm. At the moment, there are a very few studies addressed the analyzes of stock market by methods of fuzzy logic, so most likely this area will develop in the coming years. It is worth noting that the necessary tools for the task has already been created and has been implemented in several mathematical packages (fuzzy sets, linguistic variables, fuzzy inference methods fuzzification and defuzzification).

Keywords: Fuzzy logic; soft computing; linguistic variables; Mamdani's fuzzy inference method; stock market; trading strategy; algorithmic trading

Научный руководитель: **Гусин В.Б.**, кандидат физико-математических наук, профессор Департамента анализа данных, принятия решений и финансовых технологий, Финансовый университет.

Введение

В 1965 г. Лотфи Заде опубликовал основополагающую работу, в которой изложил основы теории нечетких множеств. В своей статье [1] Заде расширил классическое понятие множества, допустив, что функция принадлежности элемента множества $\mu_A(x)$ может принимать любые значения в интервале $[0,1]$, а не только четкие значения 0 или 1.

Теория, предложенная Заде, несмотря на сравнительную простоту, отлично зарекомендовала себя при решении практических задач и в настоящий момент находится в периоде своего бурного развития.

В данной статье представлен один из способов применения методов нечеткой логики при построении простейшей торговой стратегии, основанной на объеме торгов предыдущего дня и тренде цены за последние три дня.

Для удобства читателя в статье приведены необходимые определения и факты из теории нечетких множеств (см. также Волкова Е. С., Гисин В. Б. Нечеткие множества и мягкие вычисления в экономике и финансах. М.: Финансовый университет, 2015. 184 с.).

Лингвистические переменные

Лингвистическая переменная x – это переменная, принимающая значения из множества слов или словосочетаний некоторого естественного или искусственного языка. Множество допустимых значений лингвистической переменной называют терм-множеством.

Формально лингвистическая переменная задается пятеркой x, T, U, G, M , где x – имя переменной, T – терм-множество, каждый элемент которого (терм) представляется как нечеткое множество на универсальном множестве U ; G – это синтаксические правила, порождающие название новых термов с использованием таких квантификаторов, как «не», «очень» и «более-менее»; M – соответствующие семантические правила, задающие функции принадлежности нечетких термов, порожденных синтаксическими правилами [2].

Зачастую лингвистические переменные используются в упрощенном варианте с описанием вида x, T, U .

Будем считать, что для каждой лингвистической переменной x_i однозначно определено

терм-множество $T = \{\tilde{a}_{i1}, \dots, \tilde{a}_{in}\}$. Выражение вида $\mu_j(x_i)$ будем называть функцией принадлежности входа x_i нечеткому терму \tilde{a}_{ij} . Вектор лингвистических переменных $X = (x_1, \dots, x_n)$ назовем базой данных.

Нечеткая модель Мамдани

Алгоритм Мамдани позволяет обрабатывать входное четкое множество на основе базы данных лингвистических переменных и системы лингвистических правил. На выходе алгоритм дает четкое решение.

Условно алгоритм можно разбить на три основных шага:

- Ввод данных и их фаззификация (приведение к нечеткости).
 - Обработка нечетких данных и логический вывод.
 - Дефаззификация (приведение к четкости) [3].
- Рассмотрим этапы алгоритма подробнее. Как уже было отмечено, база данных лингвистических переменных имеет вид: $X = (x_1, \dots, x_n)$. Система правил содержит правила следующего вида:

ЕСЛИ $(x_1 = \tilde{a}_{1j}$ и $x_2 = \tilde{a}_{2j}$ и...и $x_n = \tilde{a}_{nj}$),

ТО $y = \tilde{d}_j, j = \overline{1, m}$,

где \tilde{d}_j – нечеткое заключение j -го правила; m – количество правил в системе правил.

Степень выполнения посылки правила j для текущего входного четкого вектора $X^* = (x_1^*, \dots, x_n^*)$ рассчитывают так:

$$\mu_j(X^*) = \mu_j(x_1^*) \wedge \dots \wedge \mu_j(x_n^*), j = \overline{1, m},$$

где \wedge – Т-норма. В условиях задачи будем рассматривать нечеткое логическое семейство Заде, т.е. $x \wedge y = \min(x, y)$.

В результате логического вывода по правилу j базы знаний получаем такое нечеткое значение выходной переменной y :

$$\tilde{d}_j^* = \text{imp}(\tilde{d}_j, \mu_j(X^*)), j = \overline{1, m},$$

где imp – импликация, которая в нечетком выводе Мамдани реализуется операцией минимума, т.е. «срезанием» функции принадлежности $\mu_{d_j}(y)$ по уровню $\mu_j(X^*)$.

Результирующее нечеткое множество \tilde{y}^* получают объединением нечетких множеств $\tilde{y}^* = \tilde{a}_1^* \cup \dots \cup \tilde{a}_m^*$, что соответствует S-конорме над функциями принадлежности, которая в семействе Заде принимает вид операции максимума:

$$\mu_{y^*}(y) = \max(\mu_{a_1^*}(y), \dots, \mu_{a_m^*}(y)).$$

Четкое значение выхода y^* , соответствующее входному вектору X^* , получается дефаззификацией нечеткого множества \tilde{y}^* . Наиболее часто применяется дефаззификация по методу центра тяжести [4].

Четкое число y^* – результат работы алгоритма Мамдани.

Построение торговой стратегии

Рассмотрим пример использования лингвистических переменных и алгоритма Мамдани для построения простой торговой стратегии на рынке акций компании Microsoft. Для анализа будет рассмотрен временной промежуток 2016/01/01–2016/11/18, причем данные за период 2016/01/01–2016/06/01 будут использоваться в качестве обучающей выборки для модели. На оставшемся промежутке будет протестирована стратегия.

Все дальнейшие расчеты выполняются в программе RStudio с использованием специализированных пакетов sets, quantmod и TTR, в которых соответственно реализована работа с нечеткими множествами, биржевыми котировками и биржевыми индикаторами.

После загрузки обучающей выборки проанализируем распределение угла наклона трехдневного

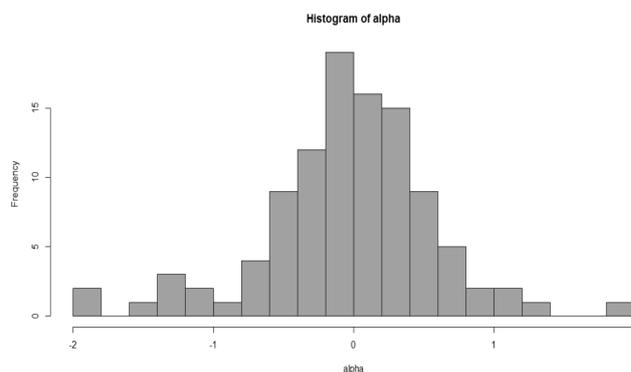


Рис. 1. Распределение угла наклона линейного тренда

тренда ($alpha$). Прямая тренда за предыдущие три дня строится методом наименьших квадратов по ценам закрытия (рис. 1).

Как видно из гистограммы, распределение угла наклона является практически симметричным относительно нуля. Исходя из этого, определим число c как квантиль уровня 0,3 распределения $alpha$, взятого по модулю и будем считать, что если $alpha \in [-c; c]$, то тренд вверх или вниз почти отсутствует. Одновременно с этим, если $alpha > 0$, фиксируется тренд вверх, если $alpha < 0$ – тренд вниз.

Причем в случае тренда вверх/вниз при увеличении по модулю $alpha$ скорость возрастания уверенности в существовании тренда должна также возрастать. Аналогично в случае отсутствия тренда скорость уверенности в его отсутствии должна возрастать при приближении $alpha$ к нулю.

На графике рис. 2 показана динамика $alpha$ и отмечен коридор отсутствия тренда.

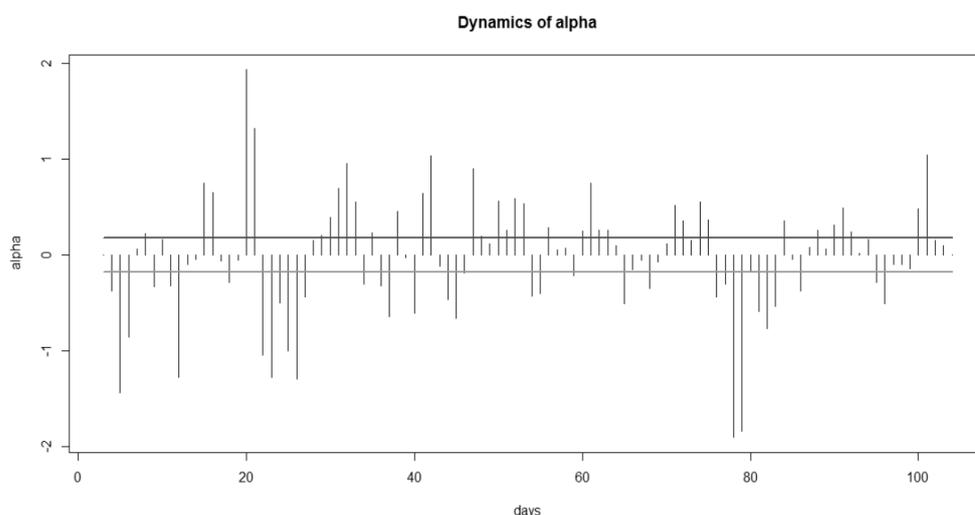


Рис. 2. Динамика alpha и коридор отсутствия тренда

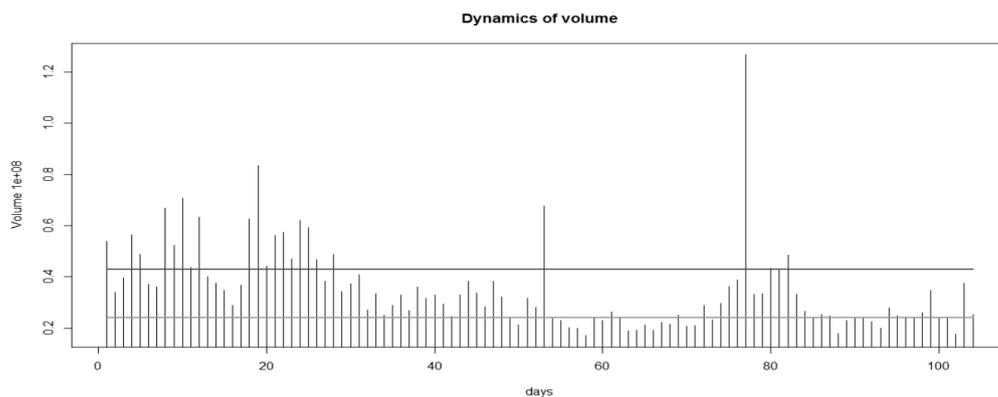


Рис. 3. Динамика volume и коридор среднего объема

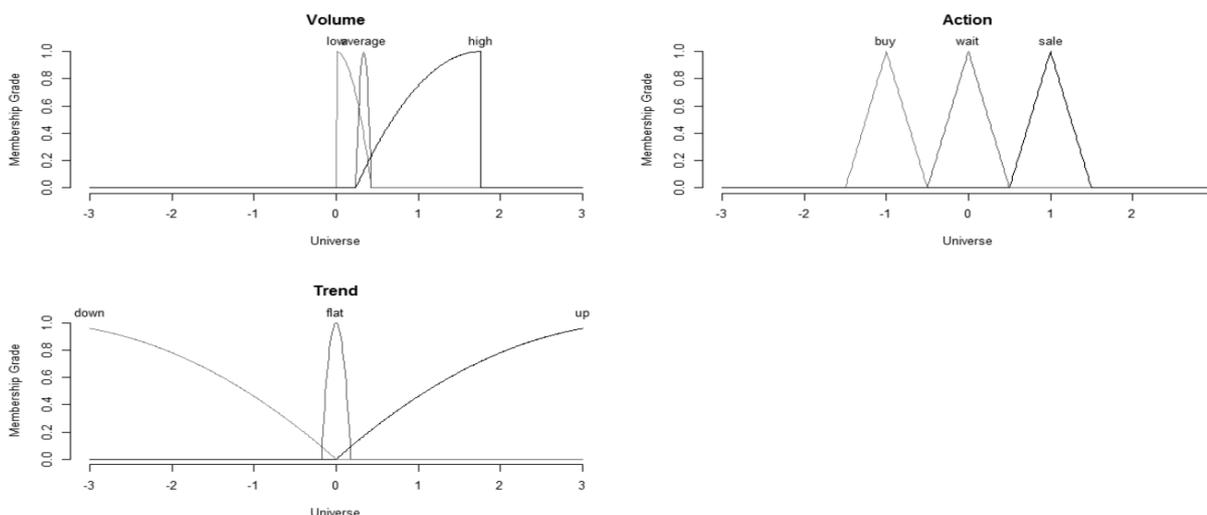


Рис. 4. Функции принадлежности лингвистических переменных

Аналогично изучим поведение объема торгов (*volume*) на обучающей выборке. Определим *a* как квантиль уровня 0,8, *b* – как квантиль уровня 0,2. Будем считать, что, если $volume < a$ – объем низкий, если $volume > b$ – объем большой, если $volume \in [b;a]$ – объем средний.

Очевидно, что при стремлении *volume* к нулю уверенность в низком объеме возрастает, при возрастании *volume* возрастает уверенность в том, что объем большой. Уверенность в том, что объем средний, возрастает при стремлении к середине отрезка $[b;a]$.

Аналогично со случаем угла наклона скорость возрастания уверенности должна увеличиваться.

На графике рис. 3 представлена динамика объема торгов и коридор среднего объема.

Завершив анализ обучающей выборки, сформируем базу Мамдани.

Правило для *T*-нормы и *S*-нормы зададим семейством нечеткой логики Заде. Общее универсальное множество определим, как $U = (-3;3)$.

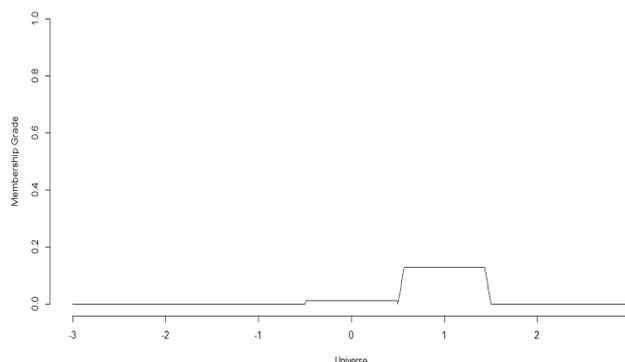


Рис. 5. Нечеткий вывод для $Volume^* = 0,25, Trend^* = -0,25$

Переменную нечеткого вывода *Action* зададим термами: buy, wait, sale. Термы соответствуют покупке акций, выходу из позиции и продаже. Каждый терм представляет собой треугольное число с центром в -1, 0, 1 соответственно и радиусом 0,5.

Лингвистические правила для действий на рынке задаются таблицей на основе экспертного мнения.

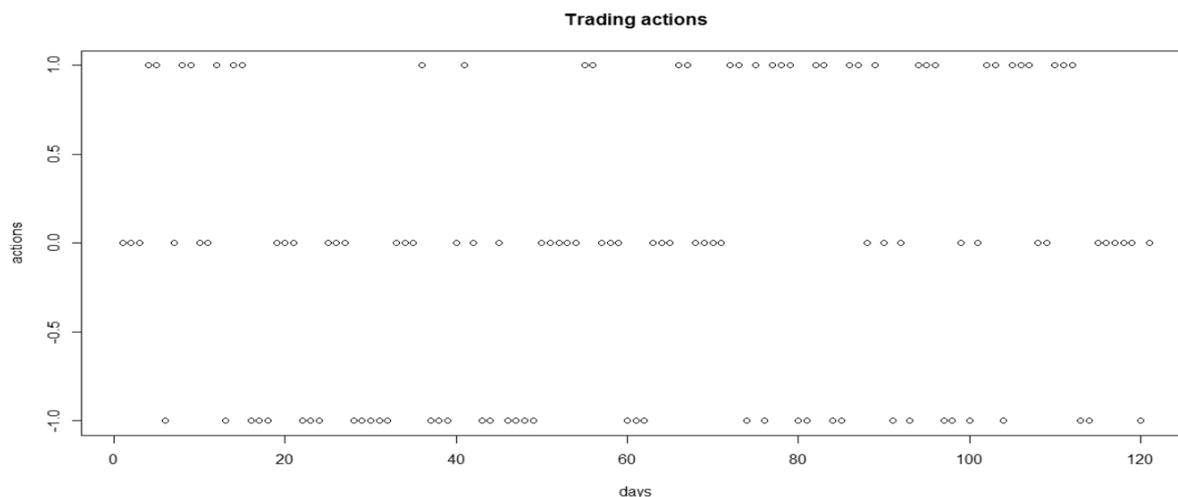


Рис. 6. Распределение торговых действий

Функции принадлежности лингвистических переменных задаются с учетом высказанных ранее предпочтений следующим образом (рис. 5) [5].

В качестве примера дадим алгоритму на вход значения $Volume^* = 0,25, Trend^* = -0,25$. В результате обработки функция принадлежности выходного нечеткого множества имеет вид, приведенный на рис. 5.

В результате дифаззификации методом центра масс получаем четкое значение $Action^* = 0,9025476$. Интерпретируем данное значение как сигнал к продаже. В дальнейшем все полученные в результате дефаззификации решения будут округляться.

Протестируем алгоритм на второй выборке 2016/06/01–2016/11/18.

Алгоритм выполнит 118 итераций. Подчеркнем, что в данном примере моделируется торговля в реальном времени без возможности получения инсайдерской информации.

В результате работы программы получено распределение торговых действий, показанное на рис. 6.

Реализация полученных торговых действий на биржевых данных дает потрясающие результаты. Так, если взглянуть на кривую доходности (equity), то в конце периода доход составляет 75,6%. Отметим, что цифры получены за чуть меньше чем полгода. Более наглядно результаты торговли представлены на рис. 7.

Если внимательно изучить кривую доходности, то оказывается, что алгоритм почти не совершал серьезных ошибок и ушел за этот период в минус только один раз, дав потерю меньше процента начальных средств.

Таблица

Правила действий на рынке

		Volume		
		low	average	high
Trend	down	buy	sale	wait
	flat	sale	buy	sale
	up	sale	buy	sale

Дополнительный анализ кривой доходности дает еще более интересные результаты (рис. 8).

За полгода работы алгоритм дал уход в минус только один раз, потеряв 0,25% от начальных средств. Стоит отметить, что это происходило в период очень низкой волатильности рынка, для которой стратегия, лежащая в основе алгоритма, плохо приспособлена.

Вместе с этим абсолютное большинство сделок были доходными, а в условиях высокой волатильности рынка кривая доходности показывала уверенный рост.

Стоит отметить, что в данной модели не учитывается комиссия брокеру и маржа от взятия акций в долг, однако ими можно пренебречь как величинами, несопоставимыми с общей доходностью.

Заключение

Было рассмотрено построение алгоритмизированной торговой стратегии, основанной на нечетком анализе данных. Данная модель показала потрясающие результаты, несмотря на простоту и очевидность предложенной лингвистической стратегии.



Рис. 7. Результат торговли с использованием алгоритма нечеткой логики

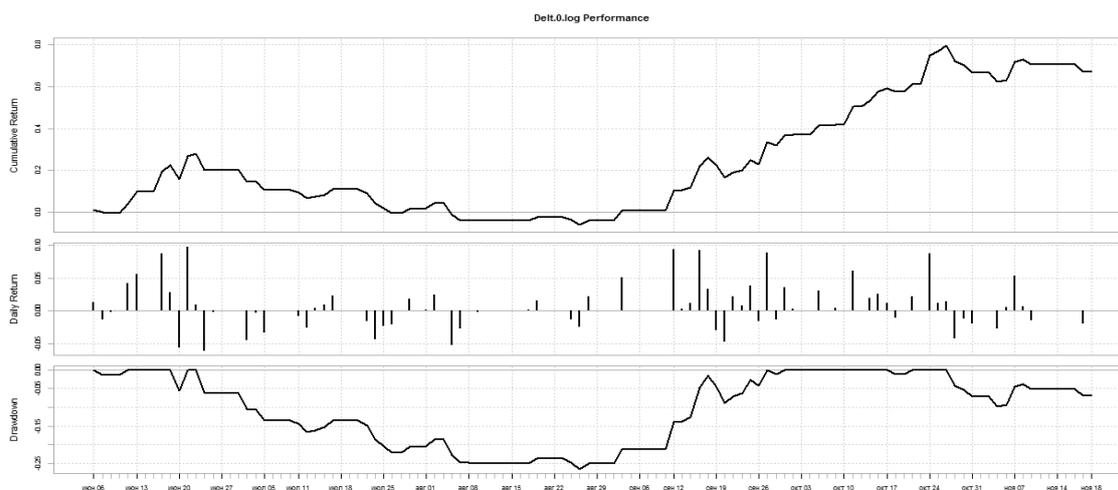


Рис. 8. Анализ проведенных сделок

На основе проделанной работы можно заключить, что применение нечетких моделей в алготрейдинге не только допустимо, но и является сверхперспективным.

Список источников

1. Zadeh L.A. Fuzzy sets // Information and control. 1965. Vol. 8. No. 3. P. 338–353.
2. Штовба С.Д. Введение в теорию нечетких множеств и нечеткую логику. URL: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php> (дата обращения: 29.09.2017).
3. Mamdani E.H. Application of fuzzy algorithms for control of simple dynamic plant // Electrical Engineers, Proceedings of the Institution of. 1974. Т. 121. № 12. С. 1585–1588.
4. Штовба С.Д. Обеспечение точности и прозрачности нечеткой модели Мамдани при обучении по экспериментальным данным // Проблемы управления и информатики. 2007. № 4. С. 102–114.
5. Heaton J. Fuzzy Logic in R // Forecasting & Futurism. 2014. № 9. С. 35–38.