ОПТИМИЗАЦИЯ ИНВЕСТИЦИОННОГО ПОРТФЕЛЯ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ НА ОСНОВЕ ПРОГНОЗИРОВАНИЯ

Юрий Зайченко, Малихех Есфандиярфард

***Аннотация**: Рассмотрена проблема оптимизации инвестиционного портфеля в условиях неопределенности с использованием прогнозирования доходностей курсов акций. Для прогнозирования курсов акций предложен нечеткий метод индуктивного моделирования- НМГУА с нечеткими входами. Приводятся результаты экспериментальных исследований- прогнозные оценки курсов акций ведущих российских компаний и полученный оптимальный портфель на основе прогнозирования и оценивается его фактическая эффективность.*

***Keywords:** fuzzy port folio optimization, stock prices forecasting, fuzzy GMDH, fuzzy inputs*

***ACM Classification Keywords**: H.4.2 Information Systems Applications: Types of Systems: Decision Support.*

***Conference**: The paper is selected from XIV^th International Conference "Knowledge-Dialogue-Solution" KDS 2008, Varna, Bulgaria, June-July 2008*

Введение

В последние годы задачи оптимизации инвестиционного портфеля представляют значительный интерес в связи с развитием финансовых рынков на Украине. Решение задачи портфельной оптимизации позволяет финансовым институтам наилучшим образом распределить имеющиеся финансовые средства в ценные бумаги (ЦБ) и уменьшить риск от ошибочных решений. Классическая задача портфельной оптимизации, впервые рассмотренная Г.Марковитцем, базируется на допущениях о нормальности распределения доходностей акций и стационарности финансовых процессов, которые на практике не выполняются.

Альтернативой классической модели Марковитца стал нечетко-множественный подход в задачах анализа и оптимизации инвестиционного портфеля, предложенный в работах [1,2] и развитый в [3]. При таком подходе доход рассматривается, как нечеткое число с функцией принадлежности треугольного и гауссовского типа, описываемое интервалами $[r_{min}, r_{cp}, r_{max}]$, где r_{min} - левый конец интервала, r_{max} - правый конец интервала, r_{cp} - наиболее вероятное значение доходности.

При этом риск портфеля трактуется как вероятность того, что реальная доходность окажется ниже некоторого критериального значения r^* (четкого ли нечеткого), задаваемого ЛПР.

Этот риск вызывается тем, что доходности акций определяются по имеющейся предыстории, т.е. основаны на прошлых данных, а реальная доходность портфеля определяется в будущий момент времени и может существенно отличаться от исходных данных. С целью снижения ожидаемого риска при построении инвестиционного портфеля целесообразно использовать не текущие, а прогнозные значения акций.

При выборе соответствующего метода прогнозирования нужно учесть специфические особенности финансовых процессов:

1) существенную нестационарность;

2) сложную неизвестную зависимость между входными и выходными переменными;

3) наличие неполной или недостоверной информации.

Одним из эффективных методов прогнозирования макроэкономики является нечеткий метод индуктивного моделирования, известный под название нечеткого МГУА, разработанный в работе [4].

В работе [5] предложен новый алгоритм нечеткого МГУА, который работает в условиях нечетких входных данных, заданных в виде интервалов неопределенности. Такая ситуация характерна для финансовых рынков, где в течении дня акции колеблются в некотором интервале.

Целью настоящей работы является исследование эффективности нечеткого МГУА с нечеткими входами для прогнозирования курсов акций с последующим использованием результатов прогнозов в задаче оптимизации инвестиционного портфеля и оценка его эффективности на реальных данных..

Вид нечеткой математической модели для треугольных ФП с нечеткими входами

Пусть имеется линейная интервальная модель регрессии:

$$Y = A_0 Z_0 + A_1 Z_1 + ... + A_n Z_n , \qquad (1)$$

где A_i – нечеткие числа треугольной формы, которые описываются тройкой параметров $A_i = (\underline{A_i}, a_i, \overline{A_i})$,

где a_i – центр интервала, $\overline{A_i}$ – его верхняя граница, $\underline{A_i}$ - нижняя граница.

В данной модели рассматривается случай симметричных функций принадлежности для параметров A_i, поэтому их можно описать парой параметров (a_i, c_i).

$\underline{A_i} = a_i - c_i$, $\overline{A_i} = a_i + c_i$, c_i – ширина интервала, c_i ≥ 0,

Z_i – также нечеткие числа треугольной формы, которые задаются параметрами $(\underline{Z_i}, \breve{Z}_i, \overline{Z_i})$, $\underline{Z_i}$ - нижняя граница, $\breve{Z}_i$ - центр, $\overline{Z_i}$ - верхняя граница нечеткого числа.

Тогда *Y* – нечеткое число, параметры которого определяются следующим образом [29]:

Центр интервала: $\breve{y} = \sum a_i * \breve{Z}_i$.

Отклонение в левой части функции принадлежности:

$$\breve{y} - \underline{y} = \sum (a_i * (\breve{Z}_i - \underline{Z}_i) + c_i \left|\breve{Z}_i\right|) . \qquad (2)$$

Откуда нижняя граница интервала: $\underline{y} = \sum (a_i * \underline{Z}_i - c_i \left|\breve{Z}_i\right|)$.

Отклонение в правой части функции принадлежности:

$$\overline{y} - \breve{y} = \sum (a_i * (\overline{Z}_i - \breve{Z}_i) + c_i \left|\breve{Z}_i\right|) = \sum a_i \overline{Z}_i - a_i \breve{Z}_i + c_i \left|\breve{Z}_i\right| .$$

Откуда верхняя границя интервала: $\overline{y} = \sum (a_i * \overline{Z}_i + c_i \left|\breve{Z}_i\right|)$.

Для того, чтобы интервальная модель была корректной, необходимо, чтобы действительное значение исходной величины *Y* принадлежало полученному в результате работы метода интервалу. Это можно описать таким образом:

$$\begin{cases} \sum (a_i * \underline{Z}_{ik} - c_i \left|\breve{Z}_{ik}\right|) \le y_k \\ \sum (a_i * \overline{Z}_{ki} + c_i \left|\breve{Z}_{ik}\right|) \ge y_k , k = \overline{1, M} \end{cases} , \qquad (3)$$

где $Z_k = \left[Z_k\right]_i$ - входная обучающая выборка, y_k - известные нам исходные значения; $k = \overline{1, M}$, *M* – количество точек наблюдения.

Следовательно, основные требования к оценочной линейной интервальной модели для треугольного частичного описания заключаются в том, чтобы найти такие значения параметров (a_i, c_i) нечетких коэффициентов, при которых:

а) наблюдаемые значения y_k попадали бы в оценочный интервал для Y_k ;

б) суммарная ширина оценочного интервала была бы минимальной.

Эти требования можно свести к задаче линейного программирования [29]:

$$\min_{a_i, c_i} \sum_{k=1}^{M} (\sum (a_i * \overline{Z}_i + c_i \left|\breve{Z}_i\right|) - \sum (a_i * \underline{Z}_i - c_i \left|\breve{Z}_i\right|)), \tag{4}$$

при условиях

$$\begin{cases} \sum (a_i * \underline{Z}_{ik} - c_i \left|\breve{Z}_{ik}\right|) \le y_k \\ \sum (a_i * \overline{Z}_{ki} + c_i \left|\breve{Z}_{ik}\right|) \ge y_k, k = \overline{1, M} \end{cases}. \tag{5}$$

Формализованная постановка задачи для треугольных ФП

Рассмотрим частичное описание вида:

$$f(x_i, x_j) = A_0 + A_1 x_i + A_2 x_j + A_3 x_i x_j + A_4 x_i^2 + A_5 x_j^2 . \tag{6}$$

Запишем его в соответствии с моделью (1). Для этого в ней нужно предположить, что $z_0 = 1$, $z_1 = x_i$, $z_2 = x_j$, $z_3 = x_i x_j$, $z_4 = x_i^2$, $z_5 = x_j^2$.

Тогда математическая модель (4) -(5) запишется в следующем виде:

$$\begin{aligned} \min_{a_i, c_i} \; & (2Mc_0 + a_1 \sum_{k=1}^{M} (\overline{x}_{ik} - \underline{x}_{ik}) + 2c_1 \sum_{k=1}^{M} \left|\breve{x}_{ik}\right| + a_2 \sum_{k=1}^{M} (\overline{x}_{jk} - \underline{x}_{jk}) + 2c_2 \sum_{k=1}^{M} \left|\breve{x}_{jk}\right| + \\ & + a_3 \sum_{k=1}^{M} (\left|\breve{x}_{ik}\right| (\overline{x}_{jk} - \underline{x}_{jk}) + \left|\breve{x}_{jk}\right| (\overline{x}_{ik} - \underline{x}_{ik})) + 2c_3 \sum_{k=1}^{M} \left|\breve{x}_{ik} \breve{x}_{jk}\right| + 2a_4 \sum_{k=1}^{M} \left|\breve{x}_{ik}\right| (\overline{x}_{ik} - \underline{x}_{ik}) + \\ & + 2c_4 \sum_{k=1}^{M} \breve{x}_{ik}^2 + 2a_5 \sum_{k=1}^{M} \left|\breve{x}_{jk}\right| (\overline{x}_{jk} - \underline{x}_{jk}) + 2c_5 \sum_{k=1}^{M} \breve{x}_{jk}^2) \end{aligned} \tag{7}$$

при условиях

$$\begin{aligned} & a_0 + a_1 \underline{x}_{ik} + a_2 \underline{x}_{jk} + a_3 (-\left|\breve{x}_{ik}\right| (\breve{x}_{jk} - \underline{x}_{jk}) - \left|\breve{x}_{jk}\right| (\breve{x}_{ik} - \underline{x}_{ik}) + \breve{x}_{ik} \breve{x}_{jk}) + \\ & + a_4 (-2\left|\breve{x}_{ik}\right| (\breve{x}_{ik} - \underline{x}_{ik}) + \breve{x}_{ik}{}^2) + a_5 (2\left|\breve{x}_{jk}\right| (\breve{x}_{jk} - \underline{x}_{jk}) + \breve{x}_{jk}{}^2) - c_0 - c_1 \left|\breve{x}_{ik}\right| - \\ & - c_2 \left|\breve{x}_{jk}\right| - c_3 \left|\breve{x}_{ik} \breve{x}_{jk}\right| - c_4 \breve{x}_{ik}^2 - c_5 \breve{x}_{jk}^2 \le y_k \\ & a_0 + a_1 \overline{x}_{ik} + a_2 \overline{x}_{jk} + a_3 (\left|\breve{x}_{ik}\right| (\overline{x}_{jk} - \breve{x}_{jk}) + \left|\breve{x}_{jk}\right| (\overline{x}_{ik} - \breve{x}_{ik}) - \breve{x}_{ik} \breve{x}_{jk}) + a_4 (2\left|\breve{x}_{ik}\right| (\overline{x}_{ik} - \\ & - \breve{x}_{ik}) - \breve{x}_{ik}{}^2) + a_5 (2\left|\breve{x}_{jk}\right| (\overline{x}_{jk} - \breve{x}_{jk}) - \breve{x}_{jk}{}^2) + c_0 + c_1 \left|\breve{x}_{ik}\right| + c_2 \left|\breve{x}_{jk}\right| + c_3 \left|\breve{x}_{ik} \breve{x}_{jk}\right| + \\ & c_4 \breve{x}_{ik}^2 + c_5 \breve{x}_{jk}^2 \ge y_k \\ & c_l \ge 0, \;\; l = \overline{0,5} . \end{aligned} \tag{8}$$

Как видно, эта задача является задачей линейного программирования, но поскольку нет ограничений неотрицательности для переменных a_i, для ее решения переходим к двойственной задаче, вводя двойственные переменные $\{\delta_k\}$ и $\{\delta_{k+M}\}$.

Решив двойственную задачу симплекс–методом и найдя оптимальные значения двойственных переменных $\{\delta_k\}, \{\delta_{k+M}\}$, найдем оптимальные значения искомых переменных c_i, a_i, $i = \overline{0,5}$, а также искомую нечеткую модель для заданного частичного описания

Описание алгоритма НМГУА

Дадим краткое описание алгоритма [1].

1. Выбор общего вида модели, которым будет описываться искомая зависимость.
2. Выбор внешних критериев оптимальности (критерия регулярности $\overline{\delta}^2$ или несмещенности $N_{см}$).
3. Выбор общего вида опорной функции (вида частичных описаний), например, линейного или квадратичного.
4. Разбиение выборки на обучающую $N_{об}$ и проверочную $N_{пров}$.
5. Присваиваем нулевые значения счетчику числа моделей k и счетчику числа рядов r.
6. Генерируем новую частичную модель f_r вида (6) на обучающей выборке. Решаем задачу ЛП (7)–(8) и находим искомые значения α_i, c_i.
7. Определяем по проверочной выборке $N_{пров}$ значение внешнего критерия ($N_{см\,k}^{(r)}$) или ($\delta_k^{(2)}(r)$).
8. $k = k + 1$. Если $k \geq C_F^2$, то $k = 0$, $r = r + 1$.
9. Вычисляем наилучшее значение критерия для моделей *r*-й итерации ($N_{см}^{(r)}$ или $\delta^{(2)}(r)$). Если $r = 1$, то переходим на шаг 6, иначе – на шаг 10.
10. Если $\left|N_{см}(r) - N_{см}(r-1)\right| \leq \varepsilon$, то переходим на шаг 11, иначе – отбираем *F* лучших моделей и положив $r = r + 1$, $k = 1$, переходим на шаг 6 и выполняем следующую *(r+1)*-ю итерацию.
11. Из *F* моделей предыдущего ряда находим по критерию регуляризации наилучшую модель.

Экспериментальные исследования НМГУА в задачах прогнозирования курсов акций

В данном эксперименте входными данными являлись цены акций 4 ведущих энергетических фирм России (03.04.2006 – 18.05.2006):

EESR – акции ОАО РАО ЕЭС России,

YUKO - акции ОАО ЮКОС,

SNGSP – привилегированые акции ОАО Сургутнефтегаз,

SNGS – обычные акции ОАО Сургутнефтегаз.

Прогнозировалась цена акций иной компании: ОАО ЛУКОЙЛ за тот же самый период

Размер выборки виборки – 32 значения.

Размер обучающей выборки – 17 значений (оптимальный розмер обучающей выборки для данного эксперимента).

Были получены следующие результаты:

1. Для треугольных функций принадлежности

а) Для нормированных входних данных:

Значения критериев для данного эксперимента составили:

СКО = 0,056481

б) Для ненормированных входных данных:

СКО = 0,914998 MAPE = 0,73%

2. Для функции принадлежности Гауса (оптимальный уровень α=0,9)

а) Для нормированных входных данных:

Значение критерия для данного эксперимента составили:

СКО = 0,030464

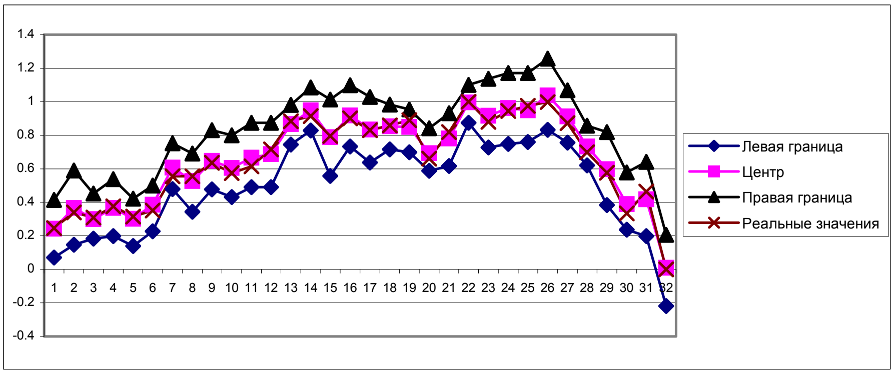

Рис. 1. Результаты эксперимента для ФП Гаусса и нормированных значений входных переменных

б) Для ненормированных входных данных:

Значения критериев для данного эксперимента составили:

СКО = 0,493511 MAPE = 0,33%

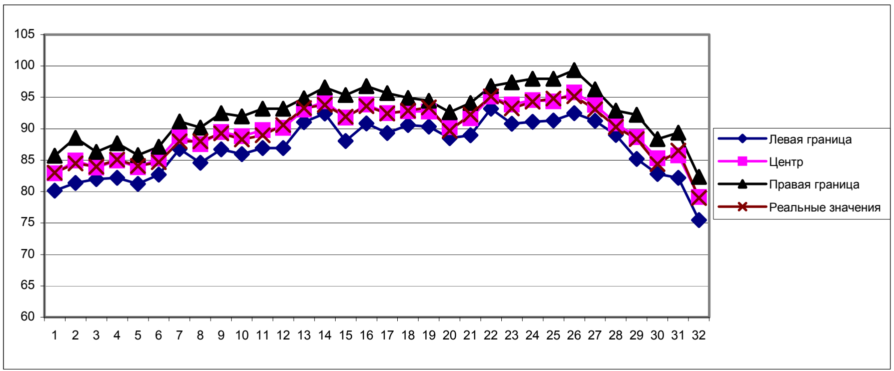

Рис. 2. Результаты эксперимента для ФП Гаусса и ненормированных значений входных переменных

Как видно из результатов экспериментов, прогнозирование с использованием треугольных и Гауссовских ФП дает хорошие результаты. При этом результаты прогнозирования с ФП Гаусса оказываются лучшими в сравнении с треугольными примерно вдвое (см. таблицу 1).

Таким образом, проведенные эксперименты подтвердили целесообразность использование НМГУА с нечеткими входными переменными для прогнозирования курсов акций.

Таблица 1. Сравнительный анализ точности прогнозирования с использованием различных видов ФП

Для нормированных данных:

	Треугольные ФП	ФП Гаусса
СКО	0,056481	0,030464

Для ненормированных данных:

	Треугольные ФП	ФП Гаусса
СКО	0,914998	0,493511
MAPE	0,73%	0,33%

Построение и анализ оптимального портфеля.

Была решена задача определения оптимального портфеля из курсов акций ведущих росийских компаний: ОАО «Лукойл»-LKOH , «Татнефть»-TATN , «Мосэнерго»-MSNG, РАО «ЕЭС»-EESR, «Газпром»- GAZP. На основе изложенного выше метода были спрогнозированы курсы акций указанных компаний и построен оптимальный портфель с таким распределением долей акций:

LKOH- 0.101, TATN -0., MSNG-0.135, EESR- 0.122, GAZP-0.642.

Для такого портфеля оценочный интервал доходности составил

$r_{min=}$ 0. 000678, r_{cp}=0.0167, r_{max}= 0.03228.

Далее были определены реальные доходности акций в следующий момент времени и для найденного оптимального портфеля подсчитана его реальная доходность. Она составила 0.014795.

Как видим, реальные данные хорошо согласуются с прогнозной доходностью и фактическая доходность попадает в оценочный интервал близко к его центру.

Выводы

1. В докладе описан метод прогнозирования курсов акций для задачи порфельной оптимизации в нечетких условиях. Для этих целей предложен нечеткий МГУА с нечеткими входами.

2. Проведены экспериментальные исследования НМГУА с нечеткими входами в задаче прогнозирования курсов акций. Результаты экспериментов подтверждают целесообразность применения нечеткого МГУА для прогнозирования курсов акций в задачах портфельной оптимизации в условиях неопределенности.

Литература

1. Недосекин А.О. Система оптимизации фондового портфеля от Сименс Бизнес Сервисез // Банковские технологии. – 2003. – № 5 – Также на сайте: http: //www. finansy. ru/publ/fin/004. htm
2. Недосекин А.О. Монотонные портфели и их оптимизация // Аудит и финансовый анализ. – 2002. – №2. - Также на сайте: http://sedok.narod.ru/s_files/PF_Article_4.zip
3. Зайченко Ю.П., Малихех Есфандиярфард. Анализ и сравнение результатов оптимизации инвестиционного портфеля при применении модели Марковица и нечетко-множественного метода // XIII-th International Conference KDS-2007. SOFIA, 2007.-Vol.1, pp.278-286.
4. Зайченко Ю. П. Нечеткий метод индуктивного моделирования в задачах прогнозирования макроэкономических показателей. // Системні дослідження та інформаційні технології.-2003.-№3.-с.-25-45.
5. Зайченко Ю.П. Нечеткий метод группового учета аргументов при неопределенных входных данных //Системні дослідження та інформаційні технології.- 2007.-№ 4.-с. 58-71.

Информация об авторах

Зайченко Юрий Петрович, *профессор, д.т.н., кафедра «Институт прикладного системного анализа». Киев, НТУУ «КПИ», ул. Политехническая 14, тел: +8(044)241-86-93, e-mail: zaych@i.com.ua*

Малихех Есфандиярфард (*Иран), аспирантка кафедры «Прикладная математика» НТУУ «КПИ», проспект Победы 37, тел: +38(096)6915857, e-mail: fard_sem@yahoo.com*