Моделирование и анализ данных
2024. Том 14. № 3. С. 87–104
DOI: https://doi.org/10.17759/mda.2024140305
ISSN: 2219-3758 (печатный)
ISSN: 2311-9454 (online)

Modelling and Data Analysis
2024. Vol. 14, no. 3, pp. 87–104
DOI: https://doi.org/10.17759/mda.2024140305
ISSN: 2219-3758 (print)
ISSN: 2311-9454 (online)

УДК 519.7

Оптимизация инвестиционного портфеля бинарным методом роя пчел

Пантелеев А.В.*
Московский авиационный институт
(национальный исследовательский университет) (МАИ)
г. Москва, Российская Федерация
ORCID: https://orcid.org/0000-0003-2493-3617
e-mail: avpanteleev@inbox.ru

Милютина С.А.**
Московский авиационный институт
(национальный исследовательский университет) (МАИ)
г. Москва, Российская Федерация
ORCID: https://orcid.org/0009-0000-5267-2157
e-mail: msofa02@mail.ru

Рассматривается задача формирования портфеля ценных бумаг, как задача бинарной оптимизации. Решение формируется с помощью разработанной модификации метода роя пчел, дополненного процедурой бинаризации с применением различных переходных функций. Исследована эффективность предложенного метода на модельных примерах и решена прикладная задача максимизации доходности портфеля с учетом ограничений на используемые средства и значение риска.

Ключевые слова: бинарная оптимизация, метаэвристические алгоритмы, метод роя пчел, переходные функции, портфель ценных бумаг.

Для цитаты:
Пантелеев А.В., Милютина С.А. Оптимизация инвестиционного портфеля бинарным методом роя пчел // Моделирование и анализ данных. 2024. Том 14. № 3. С. 87–104. DOI: https://doi.org/10.17759/mda.2024140305

***Пантелеев Андрей Владимирович,** доктор физико-математических наук, профессор, заведующий кафедрой математической кибернетики института «Компьютерные науки и прикладная математика», Московский авиационный институт (Национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация, ORCID: https://orcid.org/0000-0003-2493-3617, e-mail: avpanteleev@inbox.ru

****Милютина София Алексеевна,** выпускница бакалавриата, институт «Компьютерные науки и прикладная математика», Московский авиационный институт (Национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация, ORCID: https://orcid.org/0009-0000-5267-2157, e-mail: msofa02@mail.ru

1. ВВЕДЕНИЕ

В условиях современной экономики вопрос о распределении собственных располагаемых средств является актуальным для владельцев капитала. Инвесторы стремятся принимать решения, которые максимизируют прибыльность их портфеля, одновременно минимизируя его риск в условиях ограниченных финансовых возможностей. Классические задачи Марковица, Блэка, Тобина–Шарпа–Линтнера могут быть решены средствами нелинейного программирования [1,2]. При введении ограничений на количество активов в портфеле или вклад каждого актива в портфель математическая модель становится задачей смешанного целочисленного нелинейного квадратичного программирования, для которой требуется разработка специальных алгоритмов. Применение метаэвристических алгоритмов оптимизации позволяет эффективно исследовать пространство поиска [3] и может быть успешно использовано для решения задач оптимизации инвестиционного портфеля. Если переменные модели принимают бинарные значения, например, «включено/выключено» или «выбрано/не выбрано», становится актуальным развитие и применение методов бинарной оптимизации. Одним из подходов в их решении являются двухэтапные схемы бинаризации [4–6]. Первый этап заключается в применении передаточной функции, которая отображает значения координат решений, сгенерированных непрерывным метаэвристическим методом, в непрерывный интервал между 0 и 1. Второй этап состоит из применения правила бинаризации, которое преобразует числа в пределах этого интервала в двоичные значения. Существуют несколько типов передаточных функций, но важно отметить, что ни одна из них не превосходит другие во всех случаях их использования, что соответствует теореме No Free Lunch (NFL) [7].

В результате применения различных идей в настоящее время разработаны следующие методы бинарной оптимизации [4]: Binary Bat Algorithm, Binary Sine Cosine Algorithm, Binary Salp Swarm Algorithm, Binary Grey Wolf Optimizer, Binary Dragonfly Algorithm, Binary Whale Optimization Algorithm, Binary Magnetic Optimization Algorithm, Binary Artificial Bee Colony, Binary Firefly algorithm, Binary Flower Pollination algorithm, Binary Particle Swarm Optimization, Binary Ant Colony Optimization, Binary Bat algorithm, Binary Cat Swarm Optimization, Binary Gravitational search algorithm, Binary Harmony Search Algorithm, Binary Biogeography-Based Optimization, Binary Ant-Lion Optimizer, Binary Spotted Hyena Optimizer, Binary Emperor Penguin Optimizer, Binary Harris Hawks Optimization, Binary Equilibrium Optimizer, Binary Atom Search Optimization, Binary Jaya Algorithm, Binary Coronavirus Herd Immunity Optimizer, Binary Butterfly Optimization Algorithm, Binary Black Widow Optimization, Binary Slime Mould Algorithm, Binary Golden Eagle Optimizer, Binary Grasshopper Optimization algorithm и др.

В статье на основе оригинального алгоритма роя пчел [3, 8] с применением различных переходных функций в процессе двухэтапной бинаризации сформирован бинарный метод роя пчел. Приведен пошаговый алгоритм, решены модельные примеры и прикладная задача оптимизации портфеля ценных бумаг. По сравнению с известными модификациями [9, 10] предложенный алгоритм не уступает в эффективности.

2. ЗАДАЧА ФОРМИРОВАНИЯ ПОРТФЕЛЯ ЦЕННЫХ БУМАГ КАК ЗАДАЧА БИНАРНОЙ ОПТИМИЗАЦИИ

Введём необходимые определения.

Доходность ценной бумаги – это показатель, отражающий уровень прибыли или убытка от инвестиций в определенный актив за определенный период времени. На данный показатель влияют такие факторы, как темпы инфляции, размер выплачиваемых держателям акций дивидендов, величина налога с прибыли и т.д. Доходность рассчитывается как отношение приносимой одной акцией прибыли к ее рыночной стоимости. Доходность портфеля измеряется как средневзвешенная сумма доходностей входящих в него бумаг.

Риск ценной бумаги отражает факт, что фактическая доходность данного актива вероятно отклонится от ожидаемой доходности. Он может быть вызван различными факторами, например, изменениями в рыночных условиях, финансовым состоянием компании, экономическими и политическими событиями, стилем управления, операционной эффективностью и т.д. Риск отдельной бумаги оценивается как среднеквадратичное (стандартное) отклонение его доходности. Риск портфеля учитывает не только риски отдельных активов, входящих в портфель, но и взаимосвязи между ними, что выражается через ковариацию их доходностей. Он отражает степень возможных колебаний общей стоимости портфеля и вероятность отклонения фактической доходности от ожидаемой.

Рассматривается n инвестиционных проектов. Заданы: вектор $\overline{r} = (\overline{r}_1, \overline{r}_2, ..., \overline{r}_n)^T$, в котором $\overline{r}_i$ – средняя доходность i-го проекта; вектор $w = (w_1, w_2, ..., w_n)^T$, где w_i – размер необходимых затрат на реализацию i-го проекта; ковариационная матрица доходностей σ, где σ_{ii} – дисперсия доходности i-го проекта, а σ_{ij} – ковариация доходностей i-го и j-го проектов.

Установлены ограничения на доступные средства и допустимый уровень риска, т.е. суммарный объём финансирования не должен превышать заданного значения C, а риск собранного инвестиционного портфеля ограничен значением A.

Задача формирования инвестиционного портфеля состоит в выборе оптимального набора проектов с учетом максимизации ожидаемой доходности P при соблюдении ограничений.

Математически задача может быть описана следующим образом: требуется найти вектор $x = (x_1, x_2, ..., x_n)^T$, определённый на множестве $\{0,1\}^n$, такой что:

$$P = \sum_{i=1}^{n} \overline{r}_i \, x_i \to \max,$$

$$\sum_{i=1}^{n} w_i \, x_i \le C, \quad \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} \, x_i \, x_j} \le A,$$

$$x_i \in \{0,1\}, \quad i = 1, ..., n,$$

где $x_i = 1$, если i-й проект входит в выбранный набор, и $x_i = 0$, если – не входит.

Таким образом, задача формирования портфеля инвестиционных проектов сводится к задаче бинарной оптимизации с ограничениями типа неравенств.

3. ПОСТАНОВКА ЗАДАЧИ БИНАРНОЙ ОПТИМИЗАЦИИ

Дана целевая функция $f(x) = f(x_1, x_2, ..., x_n)$, определённая на множестве допустимых решений $D = \{0,1\}^n$.

Требуется найти глобальный условный минимум функции $f(x)$ на множестве D, т.е. найти точку $x^* \in D$, что:

$$f(x^*) = \min_{x \in D} f(x), \tag{1}$$

где $x = (x_1, x_2, ..., x_n)^T$.

Задача поиска максимума целевой функции $f(x)$ сводится к задаче поиска минимума путем замены знака перед функцией на противоположный: $f(x^*) = \max_{x \in D} f(x) = -\min_{x \in D} [-f(x)]$.

4. СТРАТЕГИЯ ПОИСКА РЕШЕНИЯ

Главная идея разработанного алгоритма решения задачи бинарной оптимизации заключается в комбинировании метаэвристического метода глобальной оптимизации с процессом бинаризации, что позволит применить его к задачам с бинарной целевой функцией.

Алгоритм основан на методе пчелиного роя, который представляет собой метаэвристический итеративный мультиагентный алгоритм [3]. Его главная идея заключается в имитации поведения пчел при поиске нектара. Для описания поведения пчёл в природе используются три следующих основных понятия.

Источник нектара характеризуется своей полезностью, которая определяется такими факторами, как удалённость от улья, концентрация нектара, удобство его добычи, а в задачах поиска глобального экстремума характеризуется значением целевой функции.

Занятые фуражиры – пчёлы, которые «связаны» с одним из источников нектара, т.е. добывают на нем нектар. Занятые фуражиры владеют следующей информацией о «своем» источнике нектара: направление от улья на источник и полезность источника.

Незанятые фуражиры – пчёлы-разведчики, которые осуществляют поиск источников нектара для их использования, а также пчёлы-наблюдатели, которые в данное время выполняют некоторые работы в улье.

Каждая незанятая пчела-фуражир может направиться к источнику нектара, следуя за пчелой-разведчиком, которая отыскала путь к этому источнику. Пчела-разведчик привлекает незанятых пчёл путем исполнения танца на специальной площадке улья, известной как «область танцев». Заинтересованные пчелы присоединяются к пчеле-разведчику и следуют за ней к месту с нектаром, становясь тем самым занятыми фуражирами.

Завершив добычу нектара, занятый фуражир возвращает его в улей и размещает там. Затем он выбирает одно из следующих действий: стать незанятым фуражиром, оставив источник нектара, продолжить сбор с текущего источника без вербовки других пчёл или выполнить вербовку. Решение о том, какое действие предпринять, принимается пчелой в соответствии с определенным вероятностным распределением.

Таким образом, на основе информации, полученной от других пчёл, пчела может:

- лететь к одному из источников нектара (двигаться к предполагаемой точке экстремума);
- оставить выбранный ранее источник и переключиться на исследование другого (выбраться из области притяжения локального экстремума);
- осуществлять новый поиск источников нектара (исследовать ранее не исследованные области).

Положение каждой пчелы задаётся вектором на множестве $\overline{D}$, причем $D \subset \overline{D}$: $x^{cont} = (x_1^{cont}, x_2^{cont}, ..., x_n^{cont})^T$, $x_i^{cont} \in \overline{D} = [-10;10], i = 1, 2, ..., n$. Используя процесс бинаризации, вектор x^{cont} преобразуется в бинарный вектор $x^{bin} = (x_1^{bin}, x_2^{bin}, ..., x_n^{bin})^T$, $x_i^{bin} \in \{0;1\}, i = 1, 2, ..., n$. Количество собранного нектара определяется значением целевой функции в точке, заданной бинарным вектором x^{bin}.

Используется двухэтапный метод бинаризации, который состоит из следующих шагов:

- передаточная функция T преобразует значения, генерируемые непрерывными методами оптимизации, в непрерывный интервал от 0 до 1;
- преобразование вещественного числа в бинарное значение по установленному правилу бинаризации.

Рассматриваются S-образные и V-образные функции, где каждая имеет по четыре варианта (рис. 1).

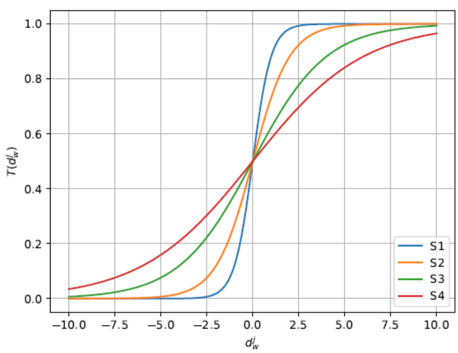

Рис. 1. а) Передаточная функция S-вида

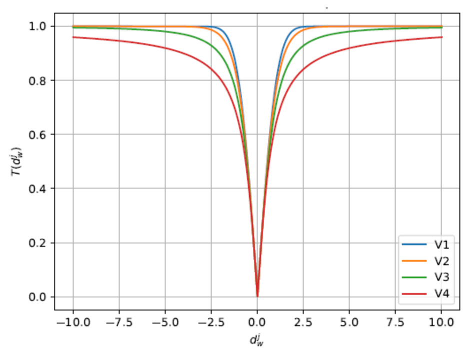

Рис. 1. б) Передаточная функция V-вида

Для того чтобы функции были временно-изменяемые, можно ввести параметр τ:

$$\tau = (1 - \frac{k}{K}) \cdot \tau_{\max} + \frac{k}{K} \cdot \tau_{\min},$$

где $\tau_{\max}$ – максимальное значение параметра τ, $\tau_{\min}$ – минимальное значение параметра τ, k – номер текущей итерации, K – максимальное число итераций. Коэффициенты $\tau_{\max}$, $\tau_{\min}$ задаются в зависимости от передаточной функции. Параметр τ изменяется с каждой итерацией, порождая новую функцию преобразования. Виды передаточных функций с введённым параметром τ представлены в табл. 1.

Таблица 1

Передаточные функции S-вида и V-вида

S-вид		V-вид	
T_{S1}	$T(x_i,\tau)=\dfrac{1}{1+e^{\frac{-2x_i}{\tau}}}$	T_{V1}	$T(x_i,\tau)=\begin{cases}1-\dfrac{2}{1+e^{\frac{-2x_i}{\tau}}}, x_i \le 0\\ \dfrac{2}{1+e^{\frac{-2x_i}{\tau}}}-1, x_i > 0\end{cases}$
T_{S2}	$T(x_i,\tau)=\dfrac{1}{1+e^{\frac{-x_i}{\tau}}}$	T_{V2}	$T(x_i,\tau)=\begin{cases}1-\dfrac{2}{1+e^{\frac{-x_i}{\tau}}}, x_i \le 0\\ \dfrac{2}{1+e^{\frac{-x_i}{\tau}}}-1, x_i > 0\end{cases}$
T_{S3}	$T(x_i,\tau)=\dfrac{1}{1+e^{\frac{-x_i}{2\tau}}}$	T_{V3}	$T(x_i,\tau)=\begin{cases}1-\dfrac{2}{1+e^{\frac{-x_i}{2\tau}}}, x_i \le 0\\ \dfrac{2}{1+e^{\frac{-x_i}{2\tau}}}-1, x_i > 0\end{cases}$
T_{S4}	$T(x_i,\tau)=\dfrac{1}{1+e^{\frac{-x_i}{3\tau}}}$	T_{V4}	$T(x_i,\tau)=\begin{cases}1-\dfrac{2}{1+e^{\frac{-x_i}{3\tau}}}, x_i \le 0\\ \dfrac{2}{1+e^{\frac{-x_i}{3\tau}}}-1, x_i > 0\end{cases}$

Второй этап процесса бинаризации заключается в преобразовании непрерывных значений из промежутка $[0;\ 1]$ в бинарные значения. Существует несколько различных правил бинаризации, таких как стандартное, дополненное, статистическое и другие. Каждое из них может иметь преимущества в зависимости от условий конкретной задачи. В статье используется стандартное правило бинаризации:

$$x_i^{bin}=\begin{cases}1, & rand[0;1]\le T(x_i^{cont},\tau), \quad i=1,\dots,n,\\ 0, & \textit{иначе,}\end{cases}$$

где n – количество переменных целевой функции.

Общая схема работы бинарного алгоритма роя пчел представлена на рис. 2. Он включает более подробное описание функций, выполняемых модулем 3.

Рис. 2. Общая схема бинарного алгоритма роя пчел

5. АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ

Шаг 1. *Инициализация алгоритма.* Задать параметры:

s	число пчёл-разведчиков;
threshold	пороговое значение расстояния между пчёлами в начальный момент;
Δ	параметр области локального поиска;
B	число пчёл, посылаемых в «наилучшие области»;
P	число пчёл, посылаемых в «перспективные области»;
K	максимальное число итераций;
b	число отбираемых наилучших значений целевой функции;
$p \le s - b$	число отбираемых перспективных значений целевой функции.
τ	параметр передаточной функции

Положить число итераций $k = 0$.

Шаг 2. *Генерирование начального роя пчел.* Сформировать начальные положения s пчёл-разведчиков при помощи равномерного распределения на множестве $\overline{D}=[-10,10]^n$. В результате получить рой пчел $\{x^{1,cont},...,x^{s,cont}\}$. Если расстояние (евклидово) между двумя пчёлами-разведчиками меньше порогового значения $threshold$, остаётся пчела с меньшим значением целевой функции. Недостающие пчёлы-разведчики генерируются заново.

Шаг 3. *Подсчет значений целевой функции для всех пчёл-разведчиков.* Координаты положения каждой j-й пчелы $x_i^{j,cont}\in\overline{D}_i, i=1,...,n; j=1,...,s$ преобразовать к бинарному значению $x_i^{j,bin}$ с помощью стандартного правила бинаризации:

$$x_i^{j,bin}=\begin{cases} 1, rand[0;1]\le T(x_i^{j,cont},\tau), i=1,...,n;\ j=1,...,s, \\ \quad 0, \qquad иначе, \end{cases} \tag{2}$$

где $rand\in[0,1]$ – число, генерируемое случайным образом, используя равномерный закон распределения, $T(x_i^{j,cont},\tau)$ – выбранная передаточная функция с параметром τ.

Если полученное бинарное значение $x_i^{j,bin}$ не удовлетворяет хотя бы одному из заданных условий, оно отбрасывается. Вместо него генерируется новое положение пчелы $x_i^{j,cont}$, затем шаг 3 повторяется.

Для каждой j-й пчелы подсчитать значение целевой функции $f(x^{j,bin})$.

Упорядочить все положения пчел (решения) по возрастанию значения целевой функции: $f(x^{1,bin})\le f(x^{2,bin})\le...\le f(x^{s,bin})$.

Шаг 4. *Выбор наилучших и перспективных значений целевой функции.* Среди значений $f(x^{1,bin}), f(x^{2,bin}),..., f(x^{s,bin})$ выбрать:

b наилучших значений;

p перспективных значений, наиболее близких к наилучшим.

Шаг 5. *Нахождение «наилучших областей» и «перспективных областей».* Выбранным на шаге 4 значениям ставятся в соответствие области локального поиска, каждая из которых представляет собой гиперкуб:

- центр области определяется координатами пчелы-разведчика в бинарной форме;
- длины сторон равны $2\Delta^k$, где $\Delta^k=\Delta\left(1-\dfrac{k}{K}\right)$.

Результатом шага 5 является нахождение b «наилучших областей» и p «перспективных областей».

Шаг 6. *Генерирование новых решений.* В каждой из «наилучших областей» случайным образом (используя равномерное распределение) генерировать B новых решений, а в каждой из «перспективных областей» генерировать P новых решений (пчела-разведчик завербовала B и P пчёл соответственно, которые стали занятыми фуражирами).

Шаг 7. *Обновление популяции.* Координаты всех найденных (*вB+pP*) точек (новых решений) преобразовать с помощью стандартного правила бинаризации (2).

Если какая-либо из найденных (*вB+pP*) точек, не удовлетворяет заданным условиям, она генерируется заново.

В каждой из ($вB+pP$) точек подсчитать значения целевой функции.

Упорядочить все точки по возрастанию значения целевой функции: $f(x^{1,bin}) \le f(x^{2,bin}) \le ... \le f(x^{bB+pP,bin})$ и выбрать s первых решений $\{x^{1,bin},...,x^{s,bin}\}$.

Шаг 8. Проверить выполнение условия окончания:

- если $k < K-1$, то положить $k = k+1$ и перейти к шагу 4;
- если $k = K-1$, завершить процесс поиска и в качестве приближенного решения выбрать $x^* = x^{1,bin}$.

Замечания.

1. Важная особенность алгоритма состоит в том, что в процессе поиска новые решения остаются в пределах множества допустимых решений. Для обеспечения этого необходимо корректировать границы «наилучших» и «перспективных» областей, учитывая границы вспомогательного множества допустимых решений. Если на шаге 5 расстояние от положения пчелы-разведчика до границы вспомогательного множества превышает значение величины Δ, тогда граница «наилучшей» или «перспективной» области должна быть скорректирована в соответствии с соответствующей границей вспомогательного множества.
2. При выборе параметров алгоритма следует учитывать, что общее число пчел в рое $s \le bB + pP$. Число наилучших областей $b \le p$, так как фактически реализуется стратегия «элитизма». Кроме того, лучше задавать $B \ge P$, так как лучшие решения должны привлекать больше пчел.
3. Ни одна из передаточных функций не превосходит остальные во всех возможных случаях, поэтому алгоритм предусматривает выбор передаточной функции. Выбирается одна из восьми функций S и V-вида, представленных в табл. 1. Величина каждой передаточной функции зависит от параметра $\tau = (1-\frac{k}{K})\cdot\tau_{\max} + \frac{k}{K}\cdot\tau_{\min}$, изменяющегося с каждой новой итерацией. Для упрощения процедуры допускается использовать постоянное значение параметра $\tau = 1$.

6. МОДЕЛЬНЫЕ ПРИМЕРЫ

На основе изложенной модификации алгоритма пчелиного роя сформирован программный модуль, реализующий поиск глобального условного минимума целевой функции, который был протестирован на нескольких функциях с известным точным решением при $n = 5$ и $n = 30$.

Для оценки эффективности работы метода для каждой тестируемой функции проводятся серии из 100 решений одной и той же задачи с одними и теми же значениями параметров. Для полученной выборки x^* вычисляются:

среднее значение отклонения полученного решения от точного: $\overline{\Delta f} = \frac{1}{100}\sum_{i=1}^{100}\Delta f_i$ где $\Delta f_i = \left| f(x^*) - f^i \right|$, $f_i -$ полученное значение целевой функции;

- среднеквадратическое отклонение $\bar{\sigma}_f = \sqrt{\overline{S}_{100}}$, где $\overline{S}_{100} = \frac{1}{99}\sum_{i=1}^{100}(\Delta f_i - \overline{\Delta f})^2$;
- наименьшее значение отклонения $\Delta f_{best} = \min\limits_i \Delta f_i$;
- количество успехов $n_{\text{усп}}$ (совпадений найденного минимума с точным решением).

Пример 1. Рассматривается тестовая функция $f_1 = \sum_{i=1}^{n} x_i^2$. Область определения функции $D = \{0, 1\}^n$, точное решение $f_{\min} = f(0,..,0) = 0$.

Выполнен анализ влияния параметров метода роя пчел на результат работы алгоритма путем последовательного изменения их значений. Каждый набор значений применялся к функции f_1 при $n = 5$ с использованием одинаковой передаточной функции, выбранной произвольно. Задавалось множество $\overline{D} = [-30;\ 30]$. Результаты работы алгоритма при различных значениях параметров метода приведены в табл. 2. Параметры метода, при которых были достигнуты лучшие решения, выделены полужирным шрифтом.

Таблица 2

Подбор параметров метода для функции f_1 при $n = 5$

Параметры метода							$\overline{\Delta f}$	$\overline{\sigma_f}$	$n_{\text{усп}}$
s	*threshold*	Δ	B	P	b	K			
20	0,001	0,85	10	10	10	100	0,27	0,1991	73
20	0,001	0,85	10	10	10	300	0,28	0,2036	78
50	0,001	0,85	10	10	10	300	0,01	0,0099	99
50	0,001	0,85	10	10	20	300	0,03	0,0294	97
50	**0,001**	**1,5**	**10**	**10**	**10**	**300**	**0**	**0**	**100**

Для функции f_1 размерности $n = 5$ независимо от выбора передаточной функции T были получены одинаково хорошие результаты: все полученные решения совпали с точным. Результаты работы метода для тестируемой функции при $n = 30$ представлены в табл. 3. Лучшие достигнутые результаты выделены полужирным шрифтом.

Таблица 3

Результаты тестирования функции f_1 при $n = 30$

T	$\overline{\Delta f}$	$\overline{\sigma_f}$	Δf_{best}	$n_{\text{усп}}$
T_{S1}	0,23	0,21	0	79
T_{S2}	**0**	**0**	**0**	**100**
T_{S3}	**0**	**0**	**0**	**100**
T_{S4}	0,05	0,04797	0	95
T_{V1}	**0**	**0**	**0**	**100**
T_{V2}	0,94	1,127676	0	40
T_{V3}	0,97	0,99909	0	39
T_{V4}	0,96	1,008484	0	39

Пример 2. Рассматривается тестовая функция $f_2 = \sum_{i=1}^{n} \left[100 \left(x_{i+1} - x_i^2 \right)^2 + \left(x_i - 1 \right)^2 \right]^2$. Область определения функции $D = \{0, 1\}^n$ Точное решение: $f_{\min} = f(1,..,1) = 0$.

Выполнен анализ влияния параметров бинарного метода роя пчел на результат работы алгоритма путем последовательного изменения их значений. Каждый набор значений применялся к функции f_2 при $n = 5$ с использованием одинаковой передаточной функции, выбранной произвольно. Задавалось множество $\overline{D} = [-10; 10]$. Результаты работы алгоритма при различных значениях параметров метода приведены в табл. 4. Параметры метода, при которых были достигнуты лучшие решения, выделены полужирным шрифтом.

Таблица 4

Подбор параметров метода для функции f_2 при $n = 5$

Параметры метода							$\overline{\Delta f}$	$\overline{\sigma_f}$	$n_{\text{усп}}$
s	*threshold*	Δ	*B*	*P*	*b*	*K*			
20	0,001	0,85	10	10	10	100	0,27	0,1991	73
20	0,001	0,85	10	10	10	300	0,28	0,2036	78
50	0,001	0,85	10	10	10	300	0,01	0,0099	99
50	0,001	0,85	10	10	20	300	0,03	0,0294	97
50	**0,001**	**1,5**	**10**	**10**	**10**	**300**	**0**	**0**	**100**

Результаты работы метода в зависимости выбора передаточной функции T представлены в табл. 5 и 6. Лучшие достигнутые результаты выделены полужирным шрифтом.

Таблица 5

Результаты тестирования функции f_2 при $n = 5$

T	$\overline{\Delta f}$	$\overline{\sigma_f}$	Δf_{best}	$n_{\text{усп}}$
T_{S1}	1,41	2,5326	0	89
T_{S2}	0,32	1,1895	0	92
T_{S3}	**0,12**	**0,4703**	**0**	**97**
T_{S4}	0,2	0,7677	0	95
T_{V1}	0,28	1,0521	0	93
T_{V2}	0,28	1,0521	0	93
T_{V3}	0,36	1,3236	0	91
T_{V4}	0,68	2,9	0	83

Таблица 6

Результаты тестирования функции f_2 при $n = 30$

T	$\overline{\Delta f}$	$\overline{\sigma_f}$	Δf_{best}	$n_{\text{усп}}$
T_{S1}	27,42	4111,296	0	82
T_{S2}	**0**	**0**	**0**	**100**
T_{S3}	1	101,01	0	99
T_{S4}	20,19	2419,347	0	84
T_{V1}	13,35	1419,52	0	88
T_{V2}	350,78	22848,09	0	2
T_{V3}	371,74	24457,6	0	1
T_{V4}	362,88	26142,49	0	1

Пример 3. Тестовая функция $f_3 = \left\{\left[\sum_{i=1}^{n} \sin^2 x_i\right] - \exp\left(\sum_{i=1}^{n} x_i^2\right)\right\} \cdot \exp\left(\sum_{i=1}^{n} \sin(\sqrt{|x_i|})\right)$ с областью определения функции $D = \{0, 1\}^n$ Точное решение: $f_{\min} = f(0,..,0) = -1$.

Выполнен анализ влияния параметров бинарного метода роя пчел на результат работы алгоритма путем последовательного изменения их значений. Каждый набор значений применялся к функции f_3 при $n = 5$ с использованием одинаковой передаточной функции, выбранной произвольно. Задавалось множество $\overline{D} = [-10; 10]$. Результаты работы алгоритма при различных значениях параметров метода приведены в табл. 7. Параметры метода, при которых были достигнуты лучшие решения, выделены полужирным шрифтом.

Таблица 7

Подбор параметров метода для функции f_3

Параметры метода							$\overline{\Delta f}$	$\overline{\sigma_f}$	$n_{\text{усп}}$
s	*threshold*	Δ	B	P	b	K			
20	0,001	0,85	10	10	10	100	0,27	0,1991	73
20	0,001	0,85	10	10	10	300	0,28	0,2036	78
50	0,001	0,85	10	10	10	300	0,01	0,0099	99
50	0,001	0,85	10	10	20	300	0,03	0,0294	97
50	**0,001**	**1,5**	**10**	**10**	**10**	**300**	**0**	**0**	**100**

Результаты работы метода в зависимости выбора передаточной функции T представлены в табл. 8 и 9. Лучшие достигнутые результаты выделены жирным шрифтом.

Таблица 8

Результаты тестирования функции f_3 при $n = 5$

T	$\overline{\Delta f}$	$\overline{\sigma_f}$	Δf_{best}	$n_{усп}$
T_{S1}	0,055124236	0,0583181267	0	95
T_{S2}	**0,03307454**	**0,035727526**	**0**	**97**
T_{S3}	0,044099388	0,0471456013	0	96
T_{S4}	0,044099388	0,0471456013	0	96
T_{V1}	0,055124235	0,0583181267	0	95
T_{V2}	0,110880428	0,111804209	0	90
T_{V3}	0,055124235	0,0583181267	0	95
T_{V4}	0,1212733	0,120196728	0	89

Таблица 9

Результаты тестирования функции f_3 при $n = 30$

T	$\overline{\Delta f}$	$\overline{\sigma_f}$	Δf_{best}	$n_{усп}$
T_{S1}	0,055124236	0,0583181267	0	95
T_{S2}	**0,03307454**	**0,035727526**	**0**	**97**
T_{S3}	0,044099388	0,0471456013	0	96
T_{S4}	0,044099388	0,0471456013	0	96
T_{V1}	0,055124235	0,0583181267	0	95
T_{V2}	0,110880428	0,111804209	0	90
T_{V3}	0,055124235	0,0583181267	0	95
T_{V4}	0,1212733	0,120196728	0	89

7. ЗАДАЧА ОПТИМИЗАЦИИ ИНВЕСТИЦИОННОГО ПОРТФЕЛЯ

Имеется набор пакетов акций крупнейших российских компаний в сфере энергетики и нефтегазовой промышленности: Газпром (GAZP), ЛУКОЙЛ (LKOH), Роснефть (ROSN), НОВАТЭК (NVTK), Газпром нефть (SIBN), РуссНефть (RNFT), Сургутнефтегаз (SNGS), Славнефть-Мегионнефтегаз (MFGS), Татнефть (TATN), Башнефть (BANE). Инвестор владеет капиталом размером $C = 10$ млн. руб., который он хочет вложить в покупку некоторых ценных бумаг из представленных.

Требуется выбрать несколько пакетов акций так, чтобы инвестор мог поместить свой капитал и добиться максимальной доходности. Риск собранного инвестиционного портфеля не должен превышать значения $A = 0,04$.

В табл. 10 представлен перечень компаний, ценные бумаги которых рассматриваются в задаче, и стоимость пакета акций каждой из них.

Таблица 10

Стоимость пакета акций компаний

i	**Компания**	w_i **(млн. руб.)**
1	Газпром (GAZP)	0,1351
2	ЛУКОЙЛ (LKOH)	7,7995
3	Роснефть (ROSN)	0,5928
4	НОВАТЭК (NVTK)	1,1614
5	Газпром нефть (SIBN)	0,7414
6	РуссНефть (RNFT)	0,2233
7	Татнефть (TATN)	0,7448
8	Башнефть (BANE)	3,0580
9	Сургутнефтегаз (SNGS)	0,0332
10	Славнефть-Мегионнефтегаз (MFGS)	0,4585

Для оценки ожидаемой доходности ценных бумаг необходимо вычислить их доходность за T одинаковых временных периодов. Рассматриваются $T = 6$ промежутков времени в один месяц в период с первого ноября 2023 г. по первое мая 2024 г. Доходность за один месяц определяется как рыночная (текущая) доходность и рассчитывается по формуле:

$$r_i^j = \frac{P_i^1 - P_i^0}{P_i^0},$$

где P_i^0 – стоимость акции i -й компании в начале периода j , P_i^1 – стоимость акции i -й компании в конце периода j .

Тогда средняя доходность ценной бумаги может быть вычислена следующим образом:

$$\overline{r_i} = \frac{1}{T} \sum\nolimits_{j=1}^{T} r_i^j .$$

Нахождение элементов ковариационной матрицы доходностей осуществляется по формуле:

$$\sigma_{ij} = \operatorname{cov}(r_i, r_j) = \frac{1}{T-1} \sum\nolimits_{k=1}^{T} (r_i^k - \overline{r_i})(r_j^k - \overline{r_j}) .$$

После обработки имеющихся данных и проведения необходимых расчётов получена информация о планируемой доходности рассматриваемых акций, представленная в табл. 11, и о ковариационной матрице доходностей σ, представленная в табл. 12.

Таблица 11

Планируемая доходность акций

i	**Компания**	$\overline{r_i}$
1	Газпром (GAZP)	-0,017
2	ЛУКОЙЛ (LKOH)	0,0089
3	Роснефть (ROSN)	0,0015
4	НОВАТЭК (NVTK)	-0,025
5	Газпром нефть (SIBN)	-0,016
6	РуссНефть (RNFT)	0,0448
7	Татнефть (TATN)	0,0161
8	Башнефть (BANE)	0,027
9	Сургутнефтегаз (SNGS)	0,0074
10	Славнефть-Мегионнефтегаз (MFGS)	-0,0008

Таблица 12

Ковариационная матрица доходностей

	GAZP	**LKOH**	**ROSN**	**NVTK**	**SIBN**	**RNFT**	**TATN**	**BANE**	**SNGS**	**MFGS**
GAZP	0,00509	0,00205	-0,00047	0,00041	-0,00014	0,00030	-0,00050	0,00489	0,00417	0,00607
LKOH		0,00237	-0,00028	-0,00013	0,00017	0,00289	-0,00152	0,00452	0,00469	0,00500
ROSN			0,00046	-0,00029	0,00043	-0,00177	0,00068	-0,00151	-0,00024	-0,00102
NVTK				0,00030	-0,00032	0,00075	-0,00017	0,00040	-0,00028	0,00054
SIBN					0,00065	-0,00179	0,00053	-0,00086	0,00073	0,00069
RNFT						0,01514	-0,00524	0,00812	0,00547	0,00545
TATN							0,00206	-0,00392	-0,00265	-0,00278
BANE								0,01130	0,00781	0,00943
SNGS									0,01022	0,01110
MFGS										0,01716

Данные всех компаний и стоимости их ценных бумаг взяты с сервиса Investing (URL: https://www.investing.com).

В результате применения бинарного метода роя пчел получено решение задачи: $x = [0,\ 0,\ 1,\ 0,\ 0,\ 1,\ 1,\ 1,\ 0,\ 0]^T$, при котором ожидаемая доходность составит $P = 0{,}0894$. Сформированный портфель потребует объём финансирования в размере 4,6189 млн. руб., а его риск составит 0,02532, что не превышает введённых ограничений $C = 10$ млн. руб., $A = 0{,}04$. Время работы программы, реализующей разработанный алгоритм, составило 16.74 сек. Опираясь на полученный результат, можно сделать вывод, что искомый поднабор пакетов ценных бумаг состоит из акций следующих компаний: Роснефть, РуссНефть, Татнефть, Башнефть. Заметим, что в процессе поиска решалась задача максимизации значения целевой функции, что обсуждалось в разд. 3, а решения, не удовлетворяющие хотя бы одному ограничению задачи (на объем финансирования или предельную величину риска), в процессе поиска отбрасывались, и процедура генерирования новых решений продолжалась.

8. ЗАКЛЮЧЕНИЕ

Рассмотрена задача оптимизации инвестиционного портфеля с заданными ограничениями на доступные средства, допустимый уровень риска и бинарность значений, которые могут принимать переменные. Для решения поставленной задачи предложен алгоритм, основанный на комбинации метаэвристического метода глобальной оптимизации с процессом бинаризации. В разработанном алгоритме используется метаэвристический метод пчелиного роя, передаточные функции вида S и V, а также стандартное правило бинаризации. Эффективность предложенного метода продемонстрирована на трех модельных примерах и прикладной задаче формирования портфеля ценных бумаг. Показано, что предложенный алгоритм способен находить решения хорошего качества при приемлемых вычислительных затратах.

Литература

1. *Бадалова А.Г., Пантелеев А.В.* Промышленный риск-менеджмент. М.: Доброе слово, 2018.
2. *Бадалова А.Г., Пантелеев А.В.* Управление рисками деятельности предприятия. М.: Вузовская книга, 2017.
3. *Пантелеев А.В., Скавинская Д.В.* Метаэвристические алгоритмы глобальной оптимизации. М.: Вузовская книга, 2019.
4. *Macedo M.* et al. Overview on binary optimization using swarm-inspired algorithms // IEEE Access. 2021. Vol. 9. P. 149814–149858.
5. *Lemus-Romani J. et al.* Binarization of Metaheuristics: Is the Transfer Function Really Important? // Biomimetics. 2023. Vol. 8. No. 5. 400.
6. *Crawford B. et al.* Q-learnheuristics: Towards data-driven balanced metaheuristics //Mathematics. 2021. Vol. 9. No. 16. 1839.
7. *Wolpert D.H., Macready W.G.* No free lunch theorems for optimization // IEEE Transactions on Evolutionary Computation. 1997. Vol. 1. No.1. P. 67–82.
8. *Karaboga D., Basturk B.* On the performance of artificial bee colony (ABC) algorithm // Applied soft computing. 2008. Vol. 8. No.1. P. 687–697.
9. *Nouioua M., Li Z., Jiang S.* New binary artificial bee colony for the 0–1 Knapsack problem // Advances in Swarm Intelligence: 9th International Conference, ICSI 2018, Shanghai, China, June 17–22, 2018, Proceedings, Part I 9. Springer International Publishing, 2018.
10. *Pampará G., Engelbrecht A.P.* Binary artificial bee colony optimization // 2011 IEEE Symposium on Swarm Intelligence. IEEE, 2011. P. 1–8.

Investment Portfolio Optimization by Binary Bee Swarm Method

Andrei V. Panteleev*
Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia
ORCID: https://orcid.org/0000-0003-2493-3617
e-mail: avpanteleev@inbox.ru

Sofia A. Milyutina**
Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia
ORCID: https://orcid.org/0009-0000-5267-2157
e-mail: msofa02@mail.ru

The problem of forming a stock portfolio is considered as a binary optimization problem. The solution is formed using the developed modification of the bee swarm method, supplemented by a binarization procedure using various transition functions. The efficiency of the proposed method is studied using model examples and the applied problem of maximizing portfolio profitability is solved taking into account constraints on the funds used and the risk value.

Keywords: binary optimization, metaheuristic algorithms, bee swarm method, transition functions, stock portfolio.

For citation:
Panteleev A.V., Milyutina S.A. Investment Portfolio Optimization by Binary Bee Swarm Method. *Modelirovanie i analiz dannykh = Modelling and Data Analysis,* 2024. Vol. 14, no. 3, pp. 87–104. DOI: https://doi.org/10.17759/mda.2024140305 (In Russ., abstr. in Engl.).

References

1. Badalova A.G., Panteleev A.V. Promyshlennyj risk-menedzhment. M.: Dobroe slovo, 2018. (In Russ.).
2. Badalova A.G., Panteleev A.V. Upravlenie riskami deyatel'nosti predpriyatiya. M.: Vuzovskaya kniga, 2017.(In Russ.).
3. Panteleev A.V., Skavinskaya D.V. Metaevristicheskie algoritmy global'noj optimizacii. M.: Vuzovskaya kniga, 2019. (In Russ.).

***Andrei V. Panteleev,** Doctor of Physics and Mathematics, Professor, Head of the Department of Mathematics and Cybernetics, Institute "Computer Science and Applied Mathematics", Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia, ORCID: https://orcid.org/0000-0003-2493-361, e-mail: avpanteleev@inbox.ru

****Sofia A. Milyutina,** Bachelor's Degree Graduate of the Institute "Computer Science and Applied Mathematics", Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia, ORCID: https://orcid.org/0009-0000-5267-2157, e-mail: msofa02@mail.ru

4. Macedo M. et al. Overview on binary optimization using swarm-inspired algorithms // IEEE Access. 2021. Vol. 9. P. 149814–149858.
5. Lemus-Romani J. et al. Binarization of Metaheuristics: Is the Transfer Function Really Important? // Biomimetics. 2023. Vol. 8. No. 5. 400.
6. Crawford B. et al. Q-learnheuristics: Towards data-driven balanced metaheuristics // Mathematics. 2021. Vol. 9. No. 16. 1839.
7. Wolpert D.H., Macready W.G. No free lunch theorems for optimization // IEEE Transactions on Evolutionary Computation. 1997. Vol. 1. No.1. P. 67–82.
8. Karaboga D., Basturk B. On the performance of artificial bee colony (ABC) algorithm // Applied soft computing. 2008. Vol. 8. No.1. P. 687–697.
9. Nouioua M., Li Z., Jiang S. New binary artificial bee colony for the 0–1 Knapsack problem // Advances in Swarm Intelligence: 9th International Conference, ICSI 2018, Shanghai, China, June 17–22, 2018, Proceedings, Part I 9. Springer International Publishing, 2018.
10. Pampará G., Engelbrecht A.P. Binary artificial bee colony optimization // 2011 IEEE Symposium on Swarm Intelligence. IEEE, 2011. P. 1–8.

Получена 13.08.2024

Принята в печать 02.09.2024

Received 13.08.2024

Accepted 02.09.2024