КОНСТРУИРАНЕ НА ОПТИМАЛЕН ИНВЕСТИЦИОНЕН ПОРТФЕЙЛ С ГЕНЕТИЧЕН АЛГОРИТЪМ

Пенка Вълкова Георгиева
Бургаски свободен университет

APPLICATION OF GENETIC ALGORITHMS IN THE PROCESS OF INVESTMENT

Penka Valkova Georgieva
Burgas Free University

Abstract: *Successful management of investment portfolio containing financial assets is a process involving proper assessment of portfolio return and risk. Effective evaluation of these characteristics is an optimization problem with a high degree of difficulty due to the large number of variables, nonlinear target function and limitations. The aim in this work is to prove that genetic algorithms, being an optimization technique in the field of artificial intelligence, can be used as a tool for solving this problem with less computational resources.*

Key words: genetic algorithm, portfolio problem, portfolio return, portfolio risk, optimization

I. ВЪВЕДЕНИЕ

Управлението на финансов инвестиционен портфейл е сложен процес, в който се разграничават няколко основни етапа, всеки от които е свързан с тежки изчислителни процедури. Както при първоначалното конструиране, така и при оперативното управление на портфейла, многократно се налага да се прави избор на една от множество възможности за комбинация на различните финансови активи, като този избор зависи от измененията на текущите цени на активите и от променящите се икономически условия.

За конструиране на портфейл традиционно се използват различни линейни и нелинейни оптимизационни алгоритми, при което възникват проблеми от различен характер. Например, в някои ситуации необходимите начални условия или не са изпълнени или не могат да бъдат проверени, поради непълнота на наличната информация. Някои от проблемите при управление на инвестиционен портфейл могат да бъдат преодолени, като се конструират всички възможни портфейли и от тях се избере този, който удовлетворява всички ограничителни условия и условията на инвеститора. Но при изчерпването на всички възможности има две важни взаимосвързани условия, за които трябва да се държи сметка: обемът на необходимите пресмятания нараства многократно и поради тази причина резултатът не може да бъде получен със стандартни изчислителни процедури и средства за разумно кратко време. [1], [8], [9], [10]

Генетичните алгоритми са средство, с помощта на което могат да се получат оптимални решения в следствие от претърсване на пространството от решения и подобряване на получените решения.[2]

II. ПОРТФЕЙЛНА ЗАДАЧА

Класическа във финансовото инвестиране е Портфейлната теория, предложената от Хари Марковиц [6], като основната концепция е, че възвращаемостите на ценните книжа са случайни величини и за тях могат да се пресметнат математическото очакване и средноквадратичното отклонение, като средноквадратичното отклонение е мярка за инвестиционния риск. Когато инвеститорът притежава повече от един финансов актив, то се приема, че е конструиран инвестиционен портфейл, като възвращаемостта и риска на такъв портфейл се пресмятат по специален начин, показан по-долу.

В модела на Марковиц са наложени специфични изисквания за инвестиционната среда. По важните от тях са:

- при вземане на решения се използват само две характеристики на активите: математическото очакване на възвращаемостта и средноквадратичното отклонение на възвращаемостта;
- очакваните възвращаемости на отделните ценни книжа имат нормално вероятностно разпределение;
- инвеститорът е несклонен към риск: за дадена очаквана възвращаемост предпочита минимум риск или за даден риск предпочита максимум очаквана възвращаемост;
- не съществуват безрискови активи;
- не се извършват къси продажби;
- няма данъци;

- инфлацията е взета предвид при определяне на очакваната възвръщаемост.

Очакваната възвращаемост за портфейл от m актива е:

$$E(R_P)=\sum_{j=1}^{m}E(r_j).x_j\ ,$$

очакваният риск на портфейла е:

$$\sigma_P^2=\sum_{j=1}^{m}\sigma_j^2.x_j^2+\sum_{j=1}^{m}\sum_{s=1,\ s\neq i}^{m}\operatorname{cov}(r_i,r_s).x_i.x_s$$

където:

$E(R_P)$ е очаквана възвращаемост на портфейла; $E(r_j)$ е очаквана възвращаемост на j-тия актив; σ_P^2 е дисперсия на възвращаемостта на портфейла; σ_j^2 е дисперсия на възвращаемостта на j-тия актив; $\operatorname{cov}(r_i,r_s)=E[(r_i-E(r_i))(r_s-E(r_s))]$ е ковариация между възвращаемостите на i-тия и s-тия активи; x_j - дял на j-тия актив, като $\sum_{j=1}^{m}x_j=1$ и $x_j\geq 0$.

Според Портфейлната теория, задачата за избор на портфейл може да бъде формулирана като оптимизационна задача върху множеството на реалните числа, с квадратна целева функция и линейни ограничения [7], по следния начин:

$$\min\ \sigma_P^2=\sum_{j=1}^{m}\sigma_j^2.x_j^2+\sum_{j=1}^{m}\sum_{s=1,\ s\neq i}^{m}\operatorname{cov}(r_i,r_s).x_i.x_s\ ,$$

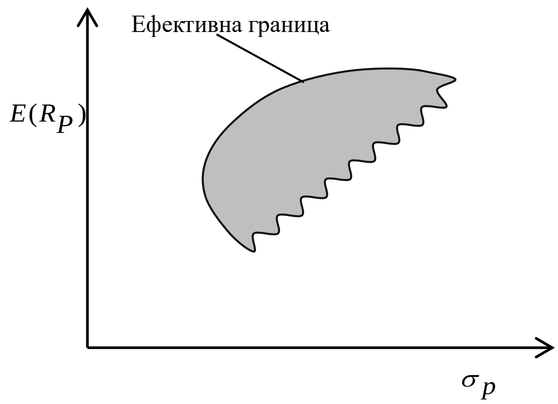
при условия:

$$E(R_P)=\sum_{j=1}^{m}E(r_j).x_j\ ,$$

$$\sum_{j=1}^{m}x_j=1,$$

$$-1\leq x_j\leq 1\qquad \text{за}\ \ j=1,2,\ldots,m\ .$$

Решенията на така формулираната оптимизационна задача формират ефективната граница, като всеки портфейл върху нея се нарича ефективен портфейл (фиг. 1).



Фигура 1. Ефективна граница.

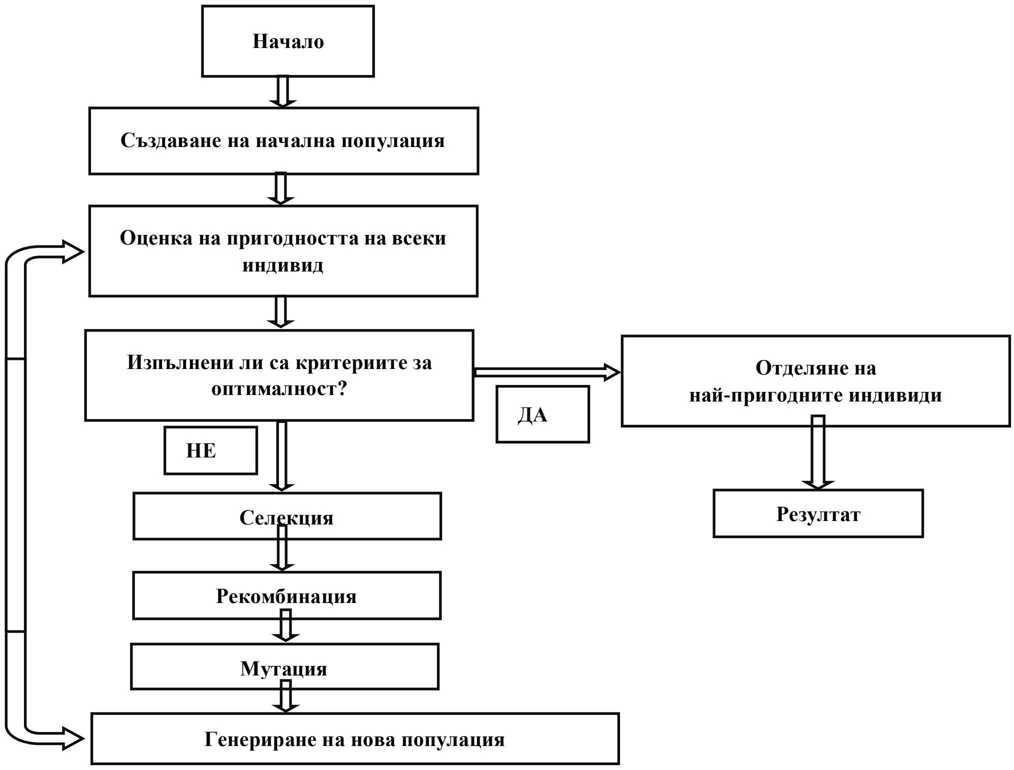
През 1990 Хари Марковиц, съвместно с Мортън Милър и Уилям Шарп, е удостоен с Нобелова награда за станалата изключително популярната теория за конструиране на инвестиционен портфейл.

III. ГЕНЕТИЧНИ АЛГОРИТМИ

Генетичните алгоритми са част от еволюционното програмиране, което е бързо разрастващата се област на изкуствения интелект. Те могат да бъдат разглеждани като алгоритми за търсене, защото при тях се изследват елементите на дадено пространство, като се използват евристични подходи, вдъхновени от природата. Всяка оптимизационна задача се представя чрез хромозоми, които са кодирано представяне на реалните стойности на променливите. Генетичните алгоритми са алгоритми за търсене, в които наподобяването на процесите в естествената еволюция се реализира чрез имитиране на принципите на генетичните изменения в природата, като целта е намиране на оптимални решения в пространството от възможни решения. За различни приложения са разработени бинарен генетичен алгоритъм, непрекъснат генетичен алгоритъм, паралелен генетичен алгоритъм, симулирано закаляване, роене, колонии от мравки и други. [4], [5]

Основните понятия на генетичните алгоритми са:

- популация – множество от възможни решения;
- хромозома – индивиди, принадлежащи на популацията, като всеки индивид е възможно решение;
- генерация – всяка популация, получена при следваща итерация;
- функция на пригодност – целева функция, която има за аргументи параметрите на хромозомата, като функционалните стойности са пропорционални на полезността или адаптивността на индивида;
- операции на генетичния алгоритъм:
  - o оценка на пригодността;
  - o селекция на популация;
  - o рекомбинация чрез кръстосване и мутация.



Фигура 2. Генетичен алгоритъм.

При изпълнението на генетичен алгоритъм се спазва следната последователност (фиг. 2):

НАЧАЛО

1. Генериране на *първоначална популация* с обем P
2. Пресмятане на стойността на функцията на пригодност за всяка хромозома
3. Създаване на *нова популация*
   - 3.1 Избор на родители чрез *селекция*
   - 3.2 Генериране на деца чрез *кръстосване*
   - 3.3 Промяна на децата чрез *мутация*
   - 3.4 Премахване на някои хромозоми от старата генерация и на децата, получени в 3.3
4. Заместване на старата популация с новополучената
5. Проверка на условие за *КРАЙ*
6. Връщане в 2.

КРАЙ

IV. РЕАЛИЗАЦИЯ

При използване на генетичен алгоритъм като оптимизационна техника, е необходимо да се дефинира целева функция със съответните ограничителни условия. В случая целта е да се намери такова разпределение на дяловете на активите в инвестиционния портфейл, че да се максимизира печалбата при най-малък възможен риск. [1], [2], [9]

Един възможен подход е да се конструира функция F на m променливи, равна на разликата на възвращаемостта и риска. Тогава при максимална възвращаемост и минимален риск, функцията F ще достига максимума си.

Така, при използване на исторически данни за цените на активите, целевата функция се дефинира по следния начин:

$$F = R_P - \sigma_P^2 = \sum_{j=1}^{m} r_j . x_j - \left(\sum_{j=1}^{m} \sigma_j^2 . x_j^2 + \sum_{j=1}^{m} \sum_{s=1,\, s\neq i}^{m} \operatorname{cov}(r_i, r_s). x_i . x_s \right),$$

където:

R_P е възвращаемостта на портфейла; r_j е възвращаемостта на j-тия актив; σ_P^2 е дисперсия на възвращаемостта на портфейла; σ_j^2 е дисперсия на възвращаемостта на j-тия актив; $\operatorname{cov}(r_i, r_s)$ е ковариация между възвращаемостите на i-тия и s-тия активи; x_j - дял на j-тия актив.

Оптимизационната задача, която се решава чрез генетичен алгоритъм, е следната:

Да се намери

$$\max F = \max\left[\sum_{j=1}^{m} r_j . x_j - \left(\sum_{j=1}^{m} \sigma_j^2 . x_j^2 + \sum_{j=1}^{m} \sum_{s=1,\, s\neq i}^{m} \operatorname{cov}(r_i, r_s). x_i . x_s \right) \right], \qquad (1)$$

при ограничителни условия

$$\sum_{j=1}^{m} x_j \leq 1, \qquad (2)$$

$$-1 \leq x_j \leq 1 \quad \text{за } j = 1, 2, \ldots, m \qquad (3)$$

Условие *(3)* предполага наличието на къси продажби на съответната фондова борса.

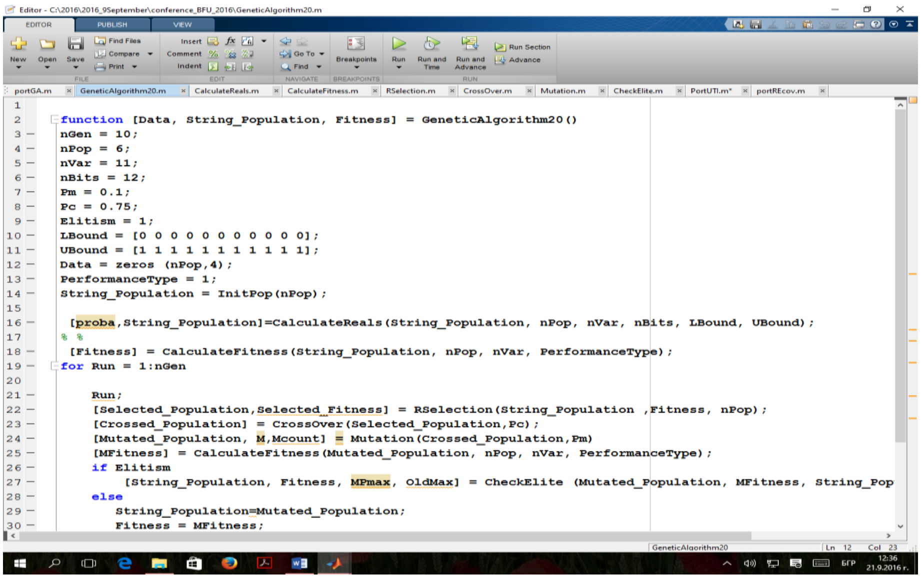
За решаване на така дефинираната оптимизационна задача са използвани данни за индекси и активи, търгувани на NYSE в периода от 15 март 2016 до 9 септември 2016: A_1=S&P 500, A_2=RUSSELL 2000 INDEX, A_3=Tesla, A_4=Apple Inc. (AAPL), A_5=Microsoft Corporation (MSFT), A_6=Cisco Systems, Inc. (CSCO), A_7=iShares Silver Trust (SLV); A_8=Halfords Group plc (HFD.L), A_9=iShares MSCI Emerging Markets (EEM), A_{10}=SPDR S&P 500 ETF (SPY), A_{11}= iPath S&P 500 VIX ST Futures ETN (VXX).

Софтуерната реализация е осъществена в средата за програмиране MatLab. Командата *PORT=GeneticAlgorithm20* стартира изпълнението на генетичния алгоритъм (фиг. 3):

function [Data, String_Population, Fitness] = GeneticAlgorithm20()

с параметри:

nGen = 10; nPop = 6; nVar = 11; nBits = 12; Pm = 0.1; Pc = 0.75; Elitism = 1;
LBound = [-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1]; UBound = [1 1 1 1 1 1 1 1 1 1 1].



Фигура 3. Част от изходния сорс-код на *GeneticAlgorithm20*

Първото поколение се генерира в

function iPopulation = InitPop(nPop, InitPop),

пресмята се пригодността на всеки индивид:

```
function Fitness = CalculateFitness(Reals_Population, nPop, nVar, PerformanceType)
Fitness = zeros(nPop, 1);
for ccount = 1:nPop
    Reals_Population(ccount, :)
        Fitness(ccount ,:)=PortUTI(Reals_Population(ccount, :), nVar, 10);
end
```

след което последователно се осъществяват селекция, рекомбинация и мутация:

```
function [Crossed_Population] = CrossOver(Selected_Population, Pc);
function [Selected_Population,Selected_Fitness] = RSelection(Population, Fitness, nPop);
function [Mutated_Population, M, Mcount] = Mutation(Population, Pm).
```

Характеристиките на активите и на портфейла се пресмятат в:

```
function [Fport] = PortUTI(x, nVars, A)
    [retur, sig, CoVa] = portREcov();
     A=0;B=0;  nVar=11;
        for j=1:nVar
            A=A+(sig(j))^2*(x(j))^2;
        end
        for j=1:nVar
            for i=1:nVar
                if i~=j
                    B=B+CoVa(i,j)*x(i)*x(j);
                end
            end
        end
    Fport= retur*x'-A-B;
```

и

```
function  [return, sig, CoVa] = portREcov()
    Price = xlsread('data_all.xlsx');
    [TT,N]=size(Price); P=Price;
            for j=1:N
              for i=1:TT
                P1(i,j)=P(TT-i+1,j);
              end
            end
            for j=1:N
                PriceP(1)=P1(1,j);
              for i=2:TT
                PriceP(i)=P1(i,j);
                rt(i)=PriceP(i)/PriceP(i-1);
```

```
            retALL(i,j)=rt(i);
          end
        end
retur=sum(retALL)/(TT-1);
        for j=1:N
          for i=2:TT
            ah(i,j)=retALL(i,j)-retur(j);
            mejd(i,j)=(ah(i,j))^2;
          end
        end
sig=sqrt(sum(mejd))/(TT-2);
CoVa=cov(retALL).
```

V. РЕЗУЛТАТИ

За тестване на програмата и проверяване на сходимостта на алгоритъма са използвани 3 различни стойности на параметъра *nGen* : 10, 100 и 1000. Получените портфейли са показани в таблица 2.

Таблица 1. Портфейли, получени при съответни стойности на *nGen*

	nGen =10	*nGen =100*	*nGen =1000*
Актив	***Дял на актива***	***Дял на актива***	***Дял на актива***
A_1	0,6476	0,8999	0,0945
A_2	-0,2874	-0,0729	-0,1888
A_3	0,9199	0,8557	0,7858
A_4	-0,5129	-0,4996	-0,3716
A_5	0,1283	0,2196	0,3934
A_6	-0,2955	-0,8339	-0,4634
A_7	0,6309	0,5209	0,7433
A_8	-0,4632	-0,3844	-0,1883
A_9	0,5225	0,6912	0,5365
A_{10}	-0,3797	-0,4794	-0,4202
A_{11}	0,0208	0	0,0435
общо	***0,9313***	***0,9171***	***0,9647***

VI. ЗАКЛЮЧЕНИЕ

Генетичните алгоритми са ефективно средство за решаване на оптимизационни задачи от различни сфери. Те търсят решения в сравнително голямо пространство от възможни решения, като получаването на решения на сложни задачи е относително бързо и надеждно.

Направените тестове показват, че проектираният и реализиран генетичен алгоритъм намира решения на портфейлната задача. Предстои да се сравнят получените решения с инвестиционни портфейли, получени с други средства – например с паралелен алгоритъм и с използване на размитата система FSSAM.

ЛИТЕРАТУРА

[1] Georgieva P., I. Popchev, St. Stoyanov. (2015). A Multi-step Procedure for Asset Allocation in Case of Limited Resources. *Cybernetics and Information Technologies*, *15*(3), 41-51.

[2] Georgieva, P. (2016). FSSAM: A Fuzzy Rule-Based System for Financial Decision Making in Real Time. *Handbook of Fuzzy Sets Comparison - Theory, Algorithms and Applications* (стр. 121-148). Science Gate Publishing P.C.

[3] Georgieva, P. V. (2016). Applying FSSAM for Currency Rates Forecasting. *Transactions on Machine Learning and Artificial Intelligence*, *4*(3), 30-40.

[4] Hoffmann F., M. Köppen., F. Klawonn, R. Roy. (2005). Soft Computing: methodologies and applications. *Series: Advances in Soft Computing, 32*.

[5] Konar, A. (2000). *Artificial Intelligence and Soft Computing Behaviour and Cognitive Modeling of Human Brain.* CRC Press.

[6] Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), 77-91.

[7] Markowitz, H. (1959.). *Portfolio Selection Efficient Diversification of Investments.* New York: John Wiley & Sons Inc.

[8] Popchev I., I. Radeva. (2006). A Decision Support Metheod for Investment Preference Evaluation. *Cibernetics and Information Technologies, 1*, 3-16.

[9] Popchev, I., P. Georgieva. (2008). A Fuzzy Approach for Solving Multicriteria Investment Problems. *Innovative Techniques in IT, E-Learning, E-assessment and Education* (стр. 427-431). New York: Springer.

[10] Radeva, I. (2013). Multi-Criteria Models for Cluster Design. *Cybernetics and Information Technologies, 13*(1), 18-33.