Application of the Improved Differential Evolution Algorithm in Portfolio

Gui-Ying NING[1,a], Dun-Qian CAO[2], and Yong-Quan ZHOU[3]

[1]*Lushan College of Guangxi University Science and Technology, Liuzhou, Guangxi 545616, China;*
[2]*College of Mathematics and Computer Science, Guangxi University for Nationalities, Nanning 530006, China;*
[3]*College of Information Science and Engineering, Guangxi University for Nationalities, Nanning, Guangxi 530006, China*

Abstract. Aiming at the NP hard problem of portfolio optimization, an improved differential evolution algorithm is proposed. In this algorithm, the mutation operator and crossover operator are set up adaptively, and then according to the characteristics of the mutation itself, two kinds of mutation operators with global search ability and local search ability are improved .The improved algorithm can improve the convergence speed and ensure the precision of the algorithm. Through five stocks of the same type and 20 different types of stocks for empirical analysis, the results show that the proposed algorithm has a certain guiding role in solving the problem of portfolio optimization.

1 Introduction

With the rapid economic development, emerging assets constantly emerging financial markets, more and more families, companies keen on capital investment, the inherent risks of financial markets and the resulting revenue has always been one of the focuses of the investors. In the market operation, anyone with an investment in its earnings also associated with certain risks at the same time, different ways of investment profit is different also. In order to find the optimal investment scheme, we need to make a trade-off between costs and benefits. The theory of portfolio comes into being.

The earliest theory of portfolio theory was proposed by American economist Markowitz [1] in 1952, which proposed mean-variance (M-V) by using two concepts of expectation and variance in mathematical statistics. Once the model is put forward, which is widely applied to portfolio selection and asset allocation. The M-V model measures the overall level of the real rate of return with the expected rate of return (mean of return), and measures the uncertainty (risk) of the rate of return by the variance (or standard deviation) of the yield. The model is:

$$\min \sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j Cov(x_i, x_j) \tag{1}$$

$$s.t \sum_{i=1}^{N} W_i \mu_i = R \tag{2}$$

[a] Corresponding author: guiyingning@126.com

$$\sum_{i=1}^{N} W_i = 1 \tag{3}$$

$$0 \le W_i \le 1, i = 1, 2, \cdots, N. \tag{4}$$

In which: W_i denotes the proportion of the funds invested in the security of i in the total investment, u_i denotes the expected return rate of the securities of i , $Cov(x_i, x_j)$ denotes the covariance between the i and the j securities, N denotes the number of total investment securities, R is the expected rate of return target. (1) represents the minimum total risk required to invest, (2) shows the expected to achieve the total revenue, (3) is the ratio of all the securities investment and is 1, (4) shows that the investment ratio of each security is between 0 and 1. The model indicates that the total risk to the investor is minimized when a certain return is satisfied.

The M-V model is established with gains and risks predictable1.

However, in actual operation process, due to the uncertainty of market factors, the risk and return of investment are often difficult to predict. Therefore, many scholars have improved the model and put forward some new models, such as semi-variance model [2], the mean absolute deviation model, such as the literatures [3-7], these methods are for the actual transaction there are some restrictions on the problem raised by the problem, which is not starting from the problem of solving the difficulty. It is proved that no matter what kind of model of portfolio optimization problem is a NP hard problem, with the continuous expansion of market investment, the traditional method can not get satisfactory results. Finding an effective solution to this problem is also of great importance.

The group intelligent algorithm is a new kind of optimization algorithm, which has achieved a very good effect in the function optimization problem. At present, some scholars have applied the algorithm to portfolio optimization problems, such as the literatures [8-12]. These methods are based on actual situation to some improvement after solving the model. The practice shows that these methods also achieved good effect. In this paper, the modified MV model is reconstructed and the improved differential evolution algorithm is used to solve it. The empirical analyses show that the proposed algorithm is effective.

2 Basic differential evolution algorithm

Differential evolution algorithm (DE) [13] was proposed by Storn and Price in 1995, which uses real number coding. The algorithm is simple and easy to implement. It has significant advantages in solving some complicated optimization problems.

2.1 The basic specific operational process of differential evolution algorithm

2.1.1 Mutation operation:

The purpose of the mutation operation is to generate a mutation vector which can disturb the target vector. The specific process as follows:

$$v_{ij}(t+1) = x_{ij}(t) + F(x_{p1,j}(t) - x_{p2,j}(t)). \tag{5}$$

In which: x_{p1} and x_{p2} are any two selected difference vector in the generation t , p_1 and p_2 are an integer of $[1, M]$,which were randomly selected random integers that are not equal to each other, F is a factor of the variation, it is usually a real number of [0,2]. x_{ij} is the jth component of the ith individual. $i = 1, 2, \cdots, M$, $j = 1, 2, \cdots, n$.

2.1.2 Crossover operation:

The purpose of crossover operation is to increase the diversity of the population. The specific operation is:

$$u_{ij}(t+1) = \begin{cases} v_{ij}(t+1) & randl_{ij} \le p_c \\ x_{ij}(t) & randl_{ij} > p_c \end{cases}. \tag{6}$$

Where p_c is the crossover probability factors, usually taking a real number of (0, 1], $randl_{ij}$ is a random number of [0, 1].

2.1.3 Selection Operation:

The selection operation is based on the evaluation of fitness; the purpose is to determine which individuals to enter the next generation. The specific operation is:

$$X_i(t+1) = \begin{cases} U_i(t+1) & if \ J(U_i(t+1)) < J(X_i(t)) \\ X_i(t) & otherwise \end{cases} \tag{7}$$

Like other evolutionary computation, the DE algorithm simulates the evolution of organisms, which makes the evolution of the population towards a better direction. The new population after mutation operation and crossover operation will choose excellent individuals into the next generation by greedy choice way .This method can guarantee the superiority of the group and accelerate the speed of searching, but also make the algorithm falls into local optimum easily. In addition, DE as a random search algorithm, the population diversity is good in the beginning, but with the increase of the number of iterations, the generation between the distributions of population density is high and the exchange of information is less and less. So as to make the global optimization ability gradually decline, despite the global optimal search can be realized, but also can appear the phenomenon of premature convergence.

3 Improvement of basic differential evolution algorithm

3.1 Parameter settings

The influence of parameter setting of differential evolution algorithm on its overall performance is very big. Variation factors F determine the deviation vector amplification ratio, the larger the F, the higher the mutation strength, the global search ability is stronger, but the slower the convergence speed. On the contrary, F is smaller, the lower the strength of individual variation, accelerate the search speed, but it may cause algorithm trapped in local optimal solution, the crossover factor CR controls whether the vector is derived from the individual or the parent of the individual variation, CR is larger, the smaller the degree of individual difference, the faster the convergence speed, but also easily caught in the local excellent. Research shows that so far $F \in [0.4,1]$ is effective, and $F = 0.5$ is usually a better choice. In order to keep the global searching ability and improve the convergence speed and precision of the algorithm, the mutation operator and the crossover operator are set as follows:

Mutation operator: $F = 1 - \dfrac{t}{t+T}$,

Crossover operator: $CR = e^{\frac{t-T}{T}}$

In which t is the current evolution generation, T is the maximum evolution generation, in this setting, mutation operator decreases in [0.5, 1] with the evolution generation increase, crossover operator exponential increases in (0, 1]. This setting can guarantee the individual diversity in the early stage of evolution and improve the global search ability. In the later stage, the degree of difference among individuals will be reduced to speed up the convergence speed.

3.2 Mutation operation settings

Differential evolution algorithm for each generation of individual x_i^t mutation to get a new individual v_i^{t+1}, commonly used mutation operations are:

1.DE/rand/1/bin, which is:
$$v_{ij}(t+1) = x_{ij}(t) + F(x_{p1,j}(t) - x_{p2,j}(t))$$

2.DE/best/2/bin, which is:
$$v_{ij}(t+1) = x_{best,j}(t) + F(x_{p1,j}(t) - x_{p2,j}(t) + x_{p3,j}(t) - x_{p4,j}(t))$$

3.DE/rand/1/bin, which is:
$$v_{ij}(t+1) = x_{i,j}(t) + F(x_{p1,j}(t) - x_{p2,j}(t) + x_{p3,j}(t) - x_{p4,j}(t))$$

The DE / rand / 1 / bin is helpful to keep the diversity of the population and improves the global searching ability, but reduces the convergence speed, the DE / best / 2 / bin reduces the differences between individuals, improves local searching ability and accelerates the convergence speed, but also accelerates the algorithm into the possibility of local optimum. Experiments show that 3.is better than 1.in enhancing the diversity of population. The 2.is very good results in improving the convergence speed and accuracy, according to the above analysis, in order to guarantee the diversity and improve the convergence speed of the algorithm, an improved mutation strategy was proposed .The new mutation operation is:

$$v_{ij}(t+1) = \lambda x_{ij}(t) + (1-\lambda)x_{best,j}(t) + F(x_{p1,j}(t) - x_{p2,j}(t) + x_{p3,j}(t) - x_{p4,j}(t)) \tag{8}$$

In which $\lambda \in [0,1]$ and $\lambda = \dfrac{T-t}{T}$, t is the current evolution generation, T is the maximum evolution generation.

In the above formula, with the increasing of evolutionary algebra t, the value of λ gradually changes from the maximum value 1 to 0 gradually. At the beginning of evolution, $\lambda = 1$,the formula becomes 3, it can guarantee the algorithm in the evolutionary initial stage has stronger global searching ability, as much as you can find the global optimal point, to the late evolution of the maximum evolutionary algebra, $\lambda = 0$,(8) becomes 2, it can guarantee the algorithm has stronger local search ability in the stage, improve the precision and convergence speed of the algorithm. The experiments results show that this setting can not only guarantee the algorithm has stronger global searching ability, and can guarantee the convergence speed and accuracy of the algorithm.

4 The application of improved differential evolution algorithm in portfolio optimization problem

In fact, the two objective functions of risks and benefits of the standard M-V model is a pair of opposite direction problem, it is not easy to solve in the process of actual operation. In order to better evaluate the contradiction between the risks and benefits. This paper considers the non-radix and the quota. A new risk weighting factor p $(0 \le p \le 1)$ is introduced, the objective functions of the risk and the benefits were combined into a target function to discuss. The combined model is as follows:

$$\min(p\sum_{i=1}^{N}\sum_{j=1}^{N}W_iW_jCov(x_i,x_j)-(1-p)\sum_{i=1}^{N}W_i\mu_i) \qquad (9)$$

$$\sum_{i=1}^{N}W_i=1 \qquad\qquad 0\leq W_i\leq 1, i=1,2,\cdots,N.$$

For a specific portfolio optimization problem, we define Eq. (9) as the objective function.

4.1 Algorithm specific process:

Step1: Initialize, randomly generates the initial population of size N , and determines the maximum evolutionary algebra T ;

Step2: Calculate the fitness value of individual species and select the best individual;

Step 3: The population is mutated according to (8);

Step 4: The crossover operation of the transformed individuals is performed according to the new crossover operator;

Step5: Calculates the fitness value of the new individual after crossing and does selection operation;

Step6: Determine whether the termination condition is reached, if so, evolution terminates, and the individual is output as the optimal solution. Otherwise, go to step7;

Step7: Sets the evolutionary algebra $t=t+1$, and into step2.

5 The empirical analysis

In order to test the effectiveness of the proposed algorithm in solving the portfolio optimization problem, we select the opening price of five Shanghai stocks, namely, Bank of China, Dongfeng Motor, China Unicom, Guodian Nanzi, Sichuan Changhong, form January to December in 2015 as the effective numerical of the test, the basic information in Table 1. In addition, select the opening price of the deep stock of four small plates, four motherboards, two gem and 10 stocks in Shanghai a total of 20 stocks from January 2015 to June 2016 in 12 months as a reference data, all the data from the Star of the Securities. In the test, we solve the optimal portfolio of each stock by calculating the monthly benefits of each stock and the covariance matrix between the two. In order to reduce the accident, each group of test data are running 10 times, the largest evolutionary algebra is 100, the population size is 10 times of the stock quantity, and take the optimal solution of 10 times as the solution result. Table 2 shows that when the risk weight factor takes different values, the fitness value and benefits / risk ratio after running 10 times, the benefits / risk ratio reflects the return on the unit risk.

Table 1. The basic information of 5 stocks

Serial number	Stock Name	Opening price (January to December)
1	China Bank(601989)	4.18; 4.24; 4.12; 4.38; 4.83; 4.34; 4.87; 4.42; 3.81; 3.79; 3.87; 4.11
2	Dongfeng Motor(600006)	5.96; 5.81; 6.21; 7.01; 8.58; 9.75; 12.48; 9.00; 13.19; 8.70; 9.52; 8.84
3	China Unicom(600050)	4.97; 4.45; 5.65; 5.46; 9.35; 8.45; 7.33; 6.12; 6.30; 6.17; 6.53; 6.23
4	Guodian Nanzi(600268)	7.35; 7.93; 8.30; 9.18; 12.12; 16.48; 14.68; 10.30; 8.55; 7.73; 9.55; 9.70
5	Sichuan Changhong(600839)	4.66; 5.69; 5.17; 5.99; 8.03; 8.52; 9.72; 7.29; 6.90; 6.10; 6.48; 6.04

Table 2. Portfolio results of five stocks

p	0	0.3	0.5	0.8	1
fitness value	-1.7897	-0.0136	-0.0050	9.1821e-005	5.7842e-004
risk	0.0204	0.0205	0.0210	0.0216	0.0268
benefits	0.0180	0.0219	0.0235	0.0246	0.0273
benefits / risk	0.8839	1.0665	1.1176	1.1401	1.0189
Optimal solution	0.2004	0.0901	0.0331	0.0365	0.1441
	0.0936	0.0717	0.0994	0.1465	0.5970
	0.5993	0.3008	0.3075	0.0974	0.0120
	0.0817	0.3781	0.2763	0.2842	0.0546
	0.0250	0.1593	0.2837	0.4354	0.1923

As we can see from the Table 2, when the risks are not taken into account, namely, $p=0$ the benefits are very little. With the increase of the risk, the benefits are also gradually increase, when only considering risk regardless of benefits, namely, $p=1$, although revenue, but earnings per unit risk declines. This shows that in the investment in stocks, income and risk is directly proportional to, so-called "high risk high reward" principle.

According to Table 2, Figure1 shows the effective frontier of the investment with different returns and corresponding minimum risk.
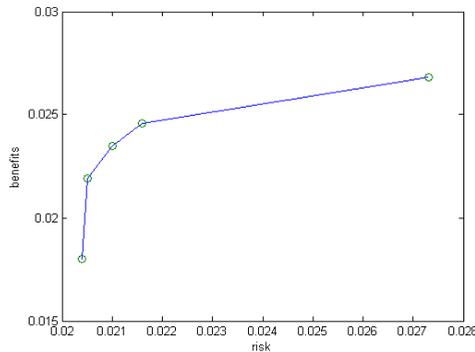

Figure 1. Effective frontier of portfolio

Table 3 presents the returns of the 20 stocks in the risk weighting factor $p=0, p=0.5, p=1$, which does not consider risk, risk and income are equally important and does not consider earnings yield per unit risk and the corresponding optimal portfolio solution.

Table 3. The results of different portfolios of 20 stocks

p	0			0.5			1		
fitness value	-0.0848			-0.0369			6.9569e-004		
risk	0.0108			0.0124			0.0187		
benefits	0.0026			0.0094			0.0112		
benefits /risk	0.2427			0.7566			0.6001		
Optimal solution	0.0438	0.0270	0.0395	0.0212	0.0001	0.0080	0.0305	0.0197	0.1322
	0.1801	0.2256	0.0001	0.0363	0.0086	0.0491	0.0329	0.0784	0.0834
	0.0247	0.0555	0.1017	0.0647	0.0022	0.0110	0.0184	0.2190	0.0477
	0.0031	0.0336	0.0825	0.2137	0.0092	0.0095	0.0879	0.0989	0.0571
	0.0004	0.0062	0.0052	0.0575	0.1795	0.0091	0.0003	0.0501	0.0084
	0.0674	0.0587	0.0361	0.0012	0.0023	0.0125	0.0044	0.0043	0.0210
	0.0012	0.0076		0.0180	0.2863		0.0038	0.0016	

Looking at the optimization results in Table 2 and Table 3, we can see that in only five small portfolio, generally high yield per unit risk, at around 1, when stock quantity increases to 20, revenues per unit risk generally dropped, which shows that the stock investment profit is associated with the number of shares, the greater the number of reduced profits, which is consistent with existing research [14]. In addition, Table 2 shows the same type of five Shanghai stocks, while table 3 is from different types of stocks. When selecting any of the five stocks in Table 3 for a combined test, we found that the return on unit risk is higher than in table 2. This shows that under the same conditions, investment stock returns are also related to the type of stock, investment in different types of stocks, the risk is relatively reduced, that is, decentralized investment is less risky than the overall investment. Figure 2 to figure 5 shows the corresponding evolutionary curves under different weights in Table 3:
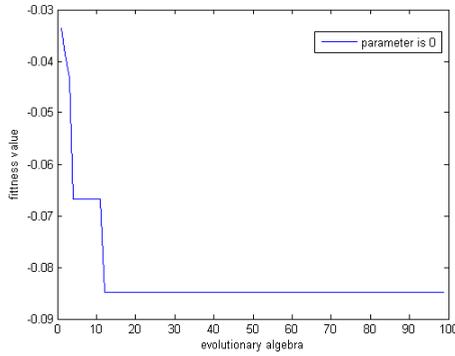

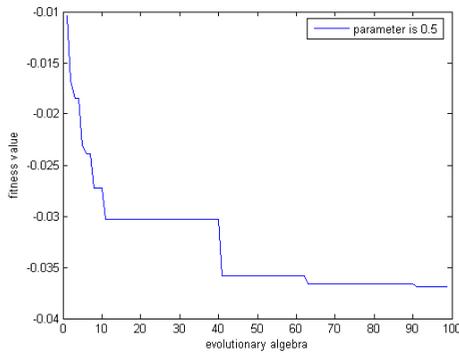**Figure 2.** Evolution curve of the risk weight factor $p = 0$

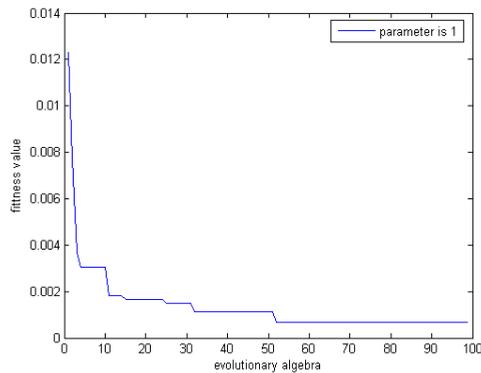**Figure 3.** Evolution curve of the risk weight factor $p = 0.5$

Figure 4. Evolution curve of the risk weight factor $p = 1$

Conclusions

Against the standard DE is slow convergence speed, easily into local optimum, this paper improves the mutation and crossover factors of standard DE, and improves the mutation operation formula. The improved DE is used to solve the NP hard problem of portfolio optimization. Through five stocks and 20 stocks for empirical testing, the test results show that the proposed algorithm can effectively guide the portfolio optimization problem.

In the actual investment portfolio, due to various factors, there will be a variety of constraints. The follow-up work of this paper is to study how to use this algorithm to solve the portfolio optimization problem under the real constraints.

Acknowledge

References

1. Markowitz H.: Portfio selection.J.J.FIN.**7** (1), 77-91 (1952).
2. Peng, Z., Zhang, W.: The positivistic multi-period mean-semi variance portfolio selection.J. Nat.Sci. S.CHN.Univ. Techn.**16** (5), 21-27 (2014).
3. Qing, Z., Wang, Z.: Robust optimization strategy of mean-semi absolute deviation model under transaction costs.J. Nat.Sci.Ludong Univ. **31**(1), 27-31 (2015).
4. Peng, Z.: The optimization on the multiperiod mean semi-absolute deviation fuzzy portfolio selection.J. F.Syst.Math.**27** (1), 168-175 (2013).
5. Jing, Y.: Mean-Variance-Approximate skewness portfolio model and empirical analysis.J. OR TRANSACTIONS.**12** (1), 107-113 (2010).
6. Peng, Z., Shu, Y.: Optimization of mean-absolute deviation fuzzy portfolio selection with transaction cost .J. Wuhan Univ. Sci.Techn. **38**(3), 235-239 (2015).
7. Lian, R., Xu, R.:.A portfolio model based on transaction costs and absolute deviation.J. Nat.Sci. Shaoyang Univ.**7** (2), 18-21 (2010).
8. Li, M., Li, Y., Fan, S.: Application of the improved artificial fish swarm algorithm in foreign exchange forecast and portfolio.J. Syst.Eng.Th. &Pract., **35**(5), 1256-1265 (2015).

9. Ma, Y., Wang, Y.: Genetic algorithm of portfolio optimization with cardinality constraint and its empirical analysis.J. Nat.Sci. N.W. Norm. Univ. **47**(2), 26-30 (2011).

10. Jiang, J., Liao, W.: Portfolio optimization based on quadratic particle swarm optimization.J. Comput. Appl.Softw.**26** (6), 161-163 (2009).

11. Li, Y., Li, P.: Research on improved model of stocks'portflio optimization based on hybrid genetic algorithm.J. Yanshan Univ.**32** (1), 65-69 (2008).

12. Du, Y., Zhou, Y.: Application of artificial glow worm swarm optimization algorithm to portfolio considering transaction cost.J. J.Comput.Appl.**34** (s2), 159-161 (2014).

13. Storn, R., Price, K.:Differential evolution a Simple and efficient heuristic for global optimization o ver continuous spaces.J. J.Glob. Optim.**11**, 341–359 (1997).

14. The editorial board of certification examination guidebook for security trading engaged persons. Sec. Inv.analys. Beijing: CHN. Comm.Publ.Co. (2002).