Multi-objective Evolutionary Algorithm Based on Data Analysis and Its Application in Portfolio Optimization

RuiWu Jiang[1(✉)] and Jing Tang[2]

[1] School of Mathematics and Statistics, Zhaotong University, Zhaotong 657000, China
ztjrw2003@tom.com
[2] Yongfeng Middle School, Zhaoyang District, Zhaotong 657000, China
ztxystxyjrw@tom.com

Abstract. The rapid development of computer technology has expanded the scope and field of use of this technology. At present, through the application of computer in optimization problems, it can effectively meet various specific objectives. Based on this, this paper discusses the multi-objective evolutionary algorithm and significance based on DEA, and analyzes the application of the algorithm in combination for reference.

Keywords: DEA model · Multi-objective evolutionary algorithm · Portfolio

1 Introduction

Because portfolio optimization problems are often NP difficult to solve, it is difficult to get the optimal solution in polynomial time, but DEA model is used as the model to evaluate at the optimal level. The portfolio optimization problem can be effectively solved by integrating it into multi-objective evolutionary algorithm. Therefore, it is necessary to analyze the algorithm and its application.

2 Discuss the DEA Based Multi-objective Evolutionary Algorithm and Its Advanced Significance

The so-called DEA, mainly refers to the production leading edge obtained by data envelopment analysis model based on known data, and the decision unit (DMU) with multi-input and multi-output is evaluated to obtain the optimal solution. At present, the DEA model mainly includes FG model, CCR model, ST model and BCC model. At present, there are many effective portfolio models which are limited by the relevant conditions to transform the problem into NP difficult solution. In general, the initial solution of intelligent optimization algorithm is random. Tabu search and other algorithms are more dependent on the initial solution, which leads to the initial search state is not

M. A. Jan and F. Khan (Eds.): BigIoT-EDU 2021, LNICST 392, pp. 454–458, 2021.
https://doi.org/10.1007/978-3-030-87903-7_55

ideal. However, the DEA model evaluation is based on the optimal level and uses the difference operator to solve the interdependence of subproblems, which can effectively avoid falling into the local optimal solution and enhance the diversity of the algorithm. Therefore, by using DEA to improve the multi-objective evolutionary algorithm, it is not necessary to determine the index weight coefficient and any form of relational expression, but also to reduce the time complexity of the algorithm and solve the problem of portfolio optimization. The specific optimization methods are as follows:

1. **The decomposition MOP, artificially decomposes it into multiple single-objective optimization problems.**

MOEA/D algorithm does not treat MOP as a whole, but as a single-objective optimization problem. When decomposing it, it mainly uses Chebyshev aggregation method. Then the optimal solution is obtained by modifying the weight vector [1]. In this case, each generation of population is picking the optimal solution and then forming the set of optimal solutions, which is different from the MOEA/D algorithm can only optimize the adjacent subproblems, which can find the optimal solution in all populations.

2. **The initial population is generated using the DEA model, which has an initial population efficiency value of**

1. DEA evaluation idea is to analyze whether the DMU input and output data are relatively effective or invalid by judging the input and output data. In the concrete operation, the input and output of the DMU should be fixed, and the DMU should be projected by mathematical planning. After that, the relative validity is evaluated according to the deviation between the DEA and the two. DEA general model is that there are n DMUs, and the input and output elements are m and s respectively$_{\mathrm{ij}}$ That means j i input to the DMU, y_{rj} The r output representing the j DMU, and the input > 0, output 0. The input weight in i is v .0_{ij}u is the r output weight$_{\mathrm{rj}}$ $X_j = \left(x_{1j}, x_{2j}, \ldots, x_{mj}\right)'$ $Y_j = \left(y_{1j}, y_{2j}, \ldots, y_{sj}\right)'$ $v = (v_1, v_2, \ldots, v_m)'$ $u = (u_1, u_2, \ldots, u_s)'$ $h_j = u'Y_j/v'X_j$ and the weight $\geq$ 0. In this case, the efficiency evaluation indicators i the evaluation decision-making unit are as follows: to satisfy the h by v and u, the appropriate choice coefficient$_{\mathrm{j}}$ $\leq$ 1. Since the CCR model is a fractional linear programming problem, it can be transformed into an equivalent linear programming problem when it is solved, and the CCR model is used when the initial population is generated.

Third, use difference operator. The whole operation of differential evolution algorithm is relatively simple and stable, and it has high optimization efficiency. After applying differential operator to multi-objective evolutionary algorithm, it can produce new individuals in the algorithm. At the same time, it can also combine the optimization algorithm based on scalar method. For the DEA model to optimize the multi-objective evolutionary algorithm, the efficiency and performance of differential operators should be improved. From the current research situation, polynomial mutation operators can be used to disturb the solution, thus increasing the local search ability of the algorithm.

Fourth, the solution that does not fully conform to the release constraint is repaired. At the same time, it is very likely that the solution that does not fully conform to the release constraint appears, so it needs to be repaired. That is, directly modify the gene bits beyond the range in the test function. At the same time, in the portfolio problem, in order to make the newly generated individual satisfied and, the last coding position of the individual should be. i when $-1 \leq i_j \leq 1 \sum_{j=1}^{N} i_j = 1 \; j_N = 1 - \sum_{j=1}^{N-1} j_j$ NWhen > 1, this means that the sum of individual values is too small. For this reason, the minimum coding position needs to be regenerated with a range of [0, 1], and then verified again until the assumed conditions are satisfied; when the i is met$_{\mathrm{N}}$ < -1, the maximum coding position needs to be regenerated in the range of [−1].

Fifth, algorithm flow and framework. 1 Population initialization. The Euclidean distance between two weight vectors is calculated, and the nearest weight T each weight vector is found, that is, the X is $i = 1, ..., N \; B(i) = \{i_1, ..., i_T\}$ i $X^{i1}, ..., X^{iT}$ T nearest weight vector is. Then, the initial population is generated in feasible space by DEA, and the reference point is initialized. Finally, the initialization EP is empty. 2 Population update. A new solution is obtained by using the difference operator to make and produce a new solution, and the polynomial is used to disturb it. Finally, it is produced, repaired and improved y, so that it is within the range of the solution. Then update the neighborhood solution and delete the dominant vector in the EP. If there is no dominant vector, add it to the EP. If the added result does not meet the stop condition, continue to update the population. If the stop condition is satisfied, stop and output the EP $B(i) k l x_k x_l y y' z F(y') F(y')$.

To sum up, the improved algorithm mainly improves the multi-objective evolutionary algorithm (MOEA/D) by DEA the initial solution, and then provides a new way for the initial solution of the algorithm. That is, the optimization of the multi-objective evolutionary algorithm with the DEA model can effectively improve the quality of the initial solution, guide the later iteration of the algorithm, and strengthen the local search ability of the whole algorithm by using the difference operator as the crossover operator [2].

3 Analysis of the Application of DEA Based Multi-objective Evolutionary Algorithm in Portfolio Optimization

To judge the optimization effect of DEA based multi-objective evolutionary algorithm in portfolio, the MOEA/D is compared with the optimized algorithm. The application is as follows.

3.1 Application of Classical M-V Models

Select the yield data for 10 stocks from 2012 to 2016, Carry out simulation experiment using Matlab. Set the initial parameter values of the MOEA/D algorithm and the DEA-MOEA/D algorithm: the target population size is 50, Variance probability and crossover probability are 0.6 and 0.5, respectively, Through random experiments on these 10 stocks, The generation distance and diversity index of each algorithm are obtained. Specific data are shown in Table 1.

Table 1. MOEA/D comparison of portfolio optimization results under DEA-MOEA/D algorithm

Mean-variance problem	Generation distance		Diversity indicators	
	Mean	Variance	Mean	Variance
MORA/D	0.0143	0.0591	0.0011	0.0003
DEA-MOEA/D	0.0115	0.0478	0.0009	0.0002

The above table shows that the optimized DEA-MOEA/D results are better than the MOEA/D algorithm, whether the generation distance or diversity index.

3.2 M-V Models with Cardinality Constraints

Still choosing yield data for 10 stocks from 2012 to 2016, Carry out simulation experiment using Matlab. Set the initial parameter values of the MOEA/D algorithm and the DEA-MOEA/D algorithm: the target population size is 50, Variance probability and crossover probability are 0.6 and 0.5, respectively, The number of iterations is 500, The scaling ratio is 0.2. Under this model, Using MOEA/D algorithm and DEA-MOEA/D algorithm to select 3, 4, 5 cardinality, Random experiments were carried out respectively. Take one of these random experiments, The results are shown in Tables 2 and 1.

Table 2. Comparison of portfolio optimization results with cardinality constraints under MOEA/D and DEA-MOEA/D algorithms

		Generation distance		Diversity indicators	
		Mean	Variance	Mean	Variance
K = 3	MOEA/D	0.0038	0.00001	0.9482	0.0258
	DEA-MOEA/D	0.0031	0.00000	0.9170	0.0314
K = 4	MOEA/D	0.0029	0.00001	0.9322	0.0517
	DEA-MOEA/D	0.0024	0.00001	0.8865	0.0840
K = 5	MOEA/D	0.0029	0.00008	0.8997	0.0511
	DEA-MOEA/D	0.0022	0.00005	0.8931	0.0416

Table 2 and Fig. 1 show that although the two algorithms are not ideal in solving the problem of cardinality constrained portfolio, the DEA-MOEA/D algorithm is closer to the front surface than the MOEA/D algorithm, but there are some solutions that do not converge. This is mainly because the function model of portfolio with cardinality constraint is discontinuous, so the result is reasonable.

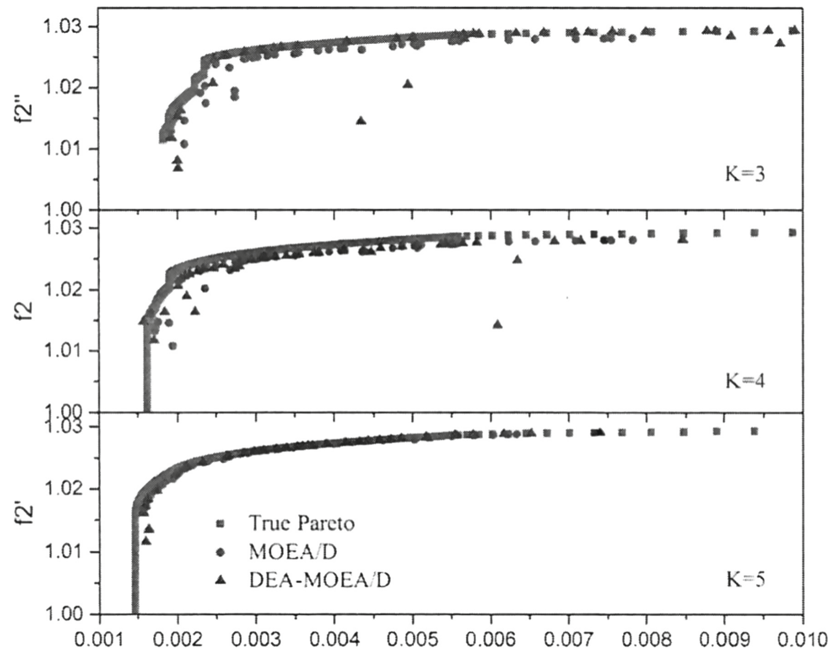

Fig. 1. Comparison of portfolio optimization results with cardinality constraints under MOEA/D and DEA-MOEA/D algorithms

Through the application of the multi-objective evolutionary algorithm based on the DEA model to the portfolio, it can be clearly seen that compared with other algorithms, the final solution of the DEA-MOEA/D algorithm is better, and the effectiveness and diversity of the front surface are enhanced. Although the application effect in portfolio with cardinality constraint is general, as shown in the figure, the performance of the algorithm will gradually increase with the increase of cardinality. All in all, by using the DEA model to optimize the multi-objective evolutionary algorithm and applying it to portfolio optimization, the optimal results can be obtained with the help of the enhanced convergence speed and the solution of diversity.

Conclusion: to sum up, DEA based multi-objective evolutionary algorithm has strong practical significance for portfolio optimization. Therefore, the DEA model should be used to optimize the multi-objective evolutionary algorithm to improve its convergence speed and increase the diversity of solutions, so as to obtain the optimal solution in the portfolio problem.

References

1. Xie, C., Long, G., Cheng, W., et al.: Advances in large-scale multi-objective evolutionary optimization algorithms. Guangxi Sci. **27**(06), 600–608 (2016)
2. Wang, Z., Li, H.: A multiobjective evolutionary algorithm using polynomial variation strategy and decomposition method. Microelectron. Comput. **38**(01), 95–100 (2021)