



An Overview of Portfolio Optimization Using Fuzzy Data

Envelopment Analysis Models

Mehrdad Rasoulzadeh*, Mohammad Fallah

Department of Industrial Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran.

PAPER INFO	ABSTRACT
Chronicle: Received: 29 April 2020 Revised: 12 July 2020 Accepted: 30 August 2020	A combination of projects, assets, programs, and other components put together in a set is called a portfolio. Arranging these components helps to facilitate the efficient management of the set and subsequently leads to achieving the strategic goals. Generally, the components of the portfolio are quantifiable and measurable which makes it possible for management to manage, prioritize, and measure different portfolios. In recent years, the portfolio in various sectors of economics, management, industry, and especially project management has been widely applied and numerous researches have been done based on mathematical models to choose the best portfolio. Among the various mathematical models, the application of data envelopment analysis models due to the unique features as well as the capability of ranking and evaluating performances has been taken by some researchers into account. In this regard, several articles have been written on selecting the best portfolio in various fields, including selecting the best stocks portfolio, selecting the best projects, portfolio of manufactured products, portfolio of patents, selecting the portfolio of assets and liabilities, etc. After presenting the Markowitz mean-variance model for portfolio optimization, these pieces of research have witnessed significant changes. Moreover, after the presentation of the fuzzy set theory by Professor Lotfizadeh, despite the ambiguities in the selection of multiple portfolios, a wide range of applications in portfolio optimization was created by combining mathematical models of portfolio optimization.
Keywords: Portfolio Optimization. Data Envelopment Analysis. Fuzzy Sets. Intuitionistic Fuzzy Sets. Markowitz Model.	

1. Introduction

Nowadays, economic and financial issues of choosing the best investment portfolio for all individuals and legal entities are of special importance and have wide aspects.



* Corresponding author

E-mail address: mehrdad.ra@gmail.com

DOI: 10.22105/jfea.2020.255034.1027

Investing means turning financial resources into one or more financial assets to achieve acceptable returns over a period of time and ultimately create wealth for the investor. Investing in monetary and capital markets is one of the most significant pillars of wealth creation and transfer in an economy [1].

Because of the prosperity of the stock market as well as the desire to invest and the influx of people's capital into this market, choosing the best investment portfolio by achieving the highest return along with the lowest risk, which means choosing the best stock portfolio, are some of the main concerns of the society. Since several factors are involved in choosing the best portfolio, including the financial structure of companies, products, sales and profitability, asset volume, political, commercial, psychological risks, etc., the selection of the right portfolio can be complicated. As a result, using traditional methods in today's complex situation will not much work. The use of mathematical models, due to their flexibility and the involvement of various quantitative and qualitative factors in the process of figuring out the best portfolio, by integrating facts such as ambiguities and doubts helps to make the answers obtained from the models closer to the facts and makes the range of multiple choices predictable. Among the various mathematical models, data envelopment analysis as a widely used technique can be effective in achieving the best portfolio. Especially when the ambiguities are applied in the model, the results of fuzzy data envelopment analysis in portfolio optimization will be much more accurate and consistent with the facts.

In the current work, first, the basic definitions of the concepts are concisely presented and then an overview of some research done in the field of portfolio optimization using fuzzy data envelopment analysis has been made.

2. Definitions

Portfolio. It is the combination of assets or plans the investor or management acquires to achieve the desired return in a given period of time, by accepting the corresponding risks, and by converting some other financial assets. This concept, especially in stock market investing, means choosing and buying a portfolio of various stocks or diverse securities aiming to increase wealth.

Data Envelopment Analysis. It is a non-parametric method in the field of operations research whose task is to evaluate the performance or efficiency of units/portfolios. This method, by considering various inputs and outputs, evaluates performances of multiple units/portfolios and identifies efficient and inefficient units/portfolios. This model was first proposed by Farrell in 1957 [10] and since then, the model has made extensive advances in various sciences and a lot of research has been done in evaluating the performance of different units as well as ranking them [2, 4].

Fuzzy. It means vague and indefinite in word. It is first introduced by Zadeh in 1965 leading to fundamental changes in the theory of classical mathematics [5]. This concept is based on logic and human decisions so that this logic can be considered as an extension of Aristotelian logic. Unlike the classical logic, which is in the state of right and wrong (zeros and ones), fuzzy logic is flexible, and the correctness or incorrectness of any logical statement is determined by the degree of membership (numerical between zero and one) [5].

In other words, instead of the characteristic function, the membership function is employed as follows:

$f_A: U \rightarrow [0,1]$.

In which $f_A(u) \in [0,1]$ refers to the level or degree of belonging of u to A .

Intuitionistic Fuzzy. Although fuzzy sets can plot and investigate the ambiguity, they cannot discuss and plot all the ambiguities that occur in real life. To this end, in the cases with insufficient information, a developed fuzzy theory called intuitive fuzzy theory is used.

An intuitive fuzzy subset of or an intuitive fuzzy subset in U is a set such as A in which to each member in $u \in U$, two degrees are assigned including membership degree and non-membership degree. In other words, for each set of A , the membership function is defined as $f_A: U \rightarrow [0,1] * [0,1]$, so that

$$F_A(u) = (\mu_A(u), \nu_A(u)), \quad 0 \leq \mu_A(u) \leq 1.$$

In the intuitive fuzzy sets, $\pi_A = 1 - \mu_A(u) - \nu_A(u)$ refers to the degree of doubt or ambiguity for each member $u \in U$.

In all of the above-mentioned definitions, $\mu_A(u)$ and $\nu_A(u)$ are both fuzzy sets. On the other hand, the values of membership in intuitive fuzzy sets can be considered as $L = \{(x, y) \in [0,1]^2 : 0 \leq x + y \leq 1\}$. Therefore, intuitive fuzzy sets are also called two-dimensional fuzzy sets [6, 7].

3. Markowitz Model

Markowitz was the first person who introduced a model capable of introducing a suitable criterion for portfolio selection by considering both risks and returns together. The Markowitz mean-variance model is the most popular selection approach for the stock portfolio. This model is based on the following statements:

- Investors are basically risk-averse and have the expected increased utility.
- Each investment option can be infinitely divisible.
- Investors select their portfolio based on the average and variance of expected returns.
- Investors have a one-period time horizon and this is the same for all investors.
- Investors prefer a higher return on a certain level of risk and, conversely, lower risk on a certain level of return.
- A commodity investment portfolio is a portfolio that has the highest return at a certain level of risk or has the lowest risk at a certain level of return [8].

4. Data Envelopment Analysis

Data envelopment analysis is a non-parametric method that is used to calculate the efficiency of homogeneous units, based on the inputs and outputs of the units, and finally the division of all units into efficient and inefficient units. This model was initially proposed by Farrell in 1957 [10] and then developed by Charnes et al. in 1978 [9] through which optimal solutions are sought by dividing a linear combination of outputs by a linear combination of inputs [9, 10].

The CCR model of data envelopment analysis is expressed as follows:

$$f_k = \max_{u,v} \frac{\sum_{n=1}^N (x_0)_{nk} v_{nk}}{\sum_{m=1}^M (x_i)_{mk} u_{mk}},$$

s. t:

$$\frac{\sum_{n=1}^N (x_0)_{nk} v_{nk}}{\sum_{m=1}^M (x_i)_{mk} u_{mk}} \leq 1 \quad j = 1, \dots, J,$$

$$v_{nk}, u_{mk} \geq 0 \quad n = 1, \dots, N, m = 1, \dots, M.$$

Since data envelopment analysis models are considered as one of the multi-criteria decision-making methods, to regard multiple indicators and criteria in selecting the best options, various types of data envelopment analysis models can be utilized. One of the applications of data envelopment analysis models is for portfolio optimization. In this regard, several types of research have been done to evaluate the efficiency and ranking of stocks as well as to identify the effective variables [10].

Due to the increasing significance of investing in the stock market, choosing the optimal portfolio is one of the most important concerns of investors for which several methods of stock portfolio selection and investment have been presented so far.

Despite the income and profit that the investor can earn from the formation of her portfolio, the issue of reducing investment risk is of particular significance. Nowadays, risk overshadows all aspects of human life and always plays an important role in all matters and decisions of individuals.

Risk conventionally is a kind of danger that is said to happen due to uncertainty about the occurrence of an accident in the future, and the higher this uncertainty is, the higher the risk [1]. There are two viewpoints to risk definition:

First Viewpoint. The risk as any possible fluctuations of economic returns in the future.

Second Viewpoint. The risk as negative fluctuations of economic returns in the future.

In all financial markets, it is the principle that investors are always risk-averse. Therefore, risk will always be present along with return as the two main pillars in portfolio selection. In 1952, Markowitz [8] for the first time, proposing the mean-variance model, proved that to form a stock portfolio, one could always minimize the risk by considering a certain level of return. Markowitz's model is known as the first mathematical model of stock portfolio optimization [2].

Before Markowitz introduced his model, it was traditionally believed that increasing diversity in the stock portfolio reduced portfolio risk, but they were unable to measure this risk. Markowitz considered the expected return per share as the average share return in previous periods and the risk per share as the variance of the return per share in previous periods. He showed that the average stock portfolio weight is equal to the stock returns, but the stock portfolio risk is not equal to the average stock weight risk. In order to calculate the expected return per share ($E(R)$), the shareholder must obtain the probable return on the

securities (R) as well as the probability of the expected return (Pr) assuming that the sum of the probabilities is equal to one. In this case, the expected return is as follows:

$$E(R) = \sum_{i=1}^m R_i Pr_i.$$

In the above formula, m represents the number of potential returns per share.

Assuming the above, the stock portfolio returns $E(R_p)$ will be equal to the weighted return of each share.

$$E(R_p) = \sum_{i=1}^m w_i E(R_i),$$

$$\sum w_i = 1.$$

Where w_i is the weight of the stock portfolio for the i th share.

In the Markowitz model, the risk per share is considered equal to the return variance ($VAR(R)$) or its second root, the standard deviation ($SD(R)$) in previous periods [3].

$$VAR(R) = \sigma^2 = \sum_{r=1}^m (R_i - E(R))^2 Pr_i,$$

$$VAR(R) = \sigma^2 = \sum_{r=1}^m (R_i - E(R))^2 Pr_i.$$

Accordingly, stock portfolio risk based on Markowitz model is shown below [2]:

$$V_{\Omega} = \text{Min} \sum_{i=1}^n w_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{\substack{l=1 \\ l \neq i}}^n w_i w_l \psi_{il}.$$

Where w_i and w_l are the weight of each share, σ_i^2 is the variance of stock returns, ψ_{il} is the covariance of double returns of shares and V_{Ω} is the variance of stock returns.

Finally, the Markowitz model with two objective functions was defined as maximizing stock portfolio return and minimizing stock portfolio risk as follows:

$$\begin{aligned} \text{Max } f_1(x) &= \sum_{i=1}^m w_i \bar{r}_i, \\ \text{Min } f_2(x) &= \sum_{i=1}^n w_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{\substack{l=1 \\ l \neq i}}^n w_i w_l \psi_{il}, \end{aligned}$$

s. t:

$$\begin{aligned} \sum w_i &= 1, \\ w_i &\geq 0. \end{aligned}$$

In financial terms, the set of Pareto optimal solutions for the portfolio selection problem is called the efficient set or the efficient boundary. In other words, for a set of assets, a set of portfolios that have the least risk for a given return is called the efficient frontier. The efficient boundary is a non-descending function that shows the best interaction between risk and return. Markowitz used the concepts of mean returns, variance, and covariance to represent the efficient boundary. This model is usually called the *EV* model, in which *E* represents the mean and *V* represents the variance [11].

Tavana et al. [3], using the Markowitz model and using 7 macro criteria and 19 indicators (8 indicators as input and 11 indicators as output), examined various stock exchange industries and chose the best industries from among them to be in the stock portfolio. After calculating the variables, he entered them into the DEA model and calculated the Relative Financial Strength Indices (RFSI) of each company using the Anderson-Peterson model. After calculating the RFSI index, companies with financial strength greater than 0.9 in the BCC-O and CCR-I models as suitable investment options were selected. After selecting the companies, the weight of each share was determined using the Markowitz model (as shown below) and solving the model through genetic technique.

$$\text{Max } \sum_{i=1}^N \mu_i x_i - \sum_{i=1}^N L(y_i) - \lambda \sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} x_i x_j,$$

S. t.:

$$\begin{aligned} \sum_{i=1}^N x_i &\leq C^0, \\ y_i &= |x_i - x_i^0| \quad i = 1, \dots, N, \\ x_i &\geq 0 \quad i = 1, \dots, N. \end{aligned}$$

Sharifghazvini et al. [12] used the Beasley index in a research to obtain a suitable mathematical model with internal market realities to extract the weights of participation of companies' stocks or the fund manager in order to obtain the variance-covariance matrix between firm returns, using the Markowitz model. In the Beasley benchmark, the average and standard deviation of each company's weekly returns and correlation coefficients between 225 Nikkei-listed companies are collected. Therefore, in order to obtain the variance-covariance matrix between firm returns, the relationship between the standard deviation of the yield of each pair of firms and the correlation coefficient between them should be used.

$$\rho_{ab} = \frac{\text{cov}(a, b)}{\sqrt{\text{var}(a) \cdot \text{var}(b)}} \Rightarrow \text{cov}(a, b) = \rho_{ab} \cdot \sqrt{\text{var}(a) \cdot \text{var}(b)}.$$

Thus, the average weekly return and the variance-covariance matrix are obtained for 225 Nikkei-listed companies in the model. In addition, because the index provided by Beasley does not specify the name of each company, information about the price and number of shares of each company is generated as a random number from the corresponding ranges in 30 selected industries of the Iranian stock market. Also in the research, in addition to reducing risk and increasing returns, minimizing portfolio costs has been added to the Markowitz model as a goal function.

$$\begin{aligned} &\text{Max} \quad \sum_{i=1}^N w_i r_i, \\ &\text{Min} \quad \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}, \\ &\text{Min} \quad \sum_{i=1}^N z_i, \\ &\text{s. t:} \end{aligned}$$

$$w_i = \frac{c_i x_i}{\sum_{i=1}^N c_i x_i} \quad i = 1, \dots, N,$$

$$\sum_{i=1}^N w_i \leq 1,$$

$$\sum_{i=1}^N v_i \leq 1,$$

$$v_i w_i \leq 0.15 \quad i = 1, \dots, N,$$

$$(1 - v_i) w_i \leq 0.1 \quad i = 1, \dots, N,$$

$$x_i \leq 0.05 S_i \quad i = 1, \dots, N,$$

$$\sum_{i \in [j]} w_i \leq 0.3 \quad j = 1, \dots, t,$$

$$\sum_{i=1}^N z_i \leq k,$$

$$x_i \in \{\mathbb{Z}\}, v_i, w_i \in \{0, 1\}, w_i \geq 0.$$

Where x_i is the number of selected stocks of type i , and v_i is a zero-one variable that specifies which company's stocks are greater than 15% of the portfolio value.

w_i is the ratio of type i shares in the portfolio and z_i is a zero and one variable to indicate the participation or non-participation of type i shares in the portfolio. Model parameters are also:

r_i : Average stock return i .

N : Variety of stocks from which the portfolio is selected.

σ_{ij} : Covariance between stocks i and j .

c_i : Stock price i .

s_i : Number of shares of the company i .

t : Number of industry.

k : Maximum portfolio variety.

In order to optimize the stock portfolio, Ahmadi et al. [13] used a combination of data envelopment analysis and heuristic factor analysis methods. In the presented research, each company is considered as a decision-making unit and input and output indicators are defined for each company. Since each of the indicators shows different dimensions of the companies' performance, it divided the output indicators into the input indicators so that the target indicators become a single and comparable indicator called performance. Finally, in order to reduce the dimensions of the problem and eliminate the correlation between the data, exploratory factor analysis was used.

Factor analysis is a set of different mathematical and statistical techniques that aim to simplify complex data sets. The main question is the answer to the question of whether a set of variables can be described in terms of the number of indicators or fewer factors than the variables and what attribute or feature each of the indicators (factors) represents. Factor analysis is used for correlation between variables.

Due to the fact that the model obtained in this research is an integer programming type, it cannot be solved by mathematical methods. Therefore, to solve the obtained model, the method of genetic algorithm and simulated annealing has been used.

5. Fuzzy Markowitz Model

Mashayekhi and Omrani [14] presented a new multi-objective model for portfolio selection including cross-performance data envelopment analysis and Markowitz mean-variance model in addition to presenting the risk and performance of the portfolio. To take the uncertainty into account, they considered the return on assets as trapezoidal fuzzy numbers and finally solved the model by employing the second type of genetic algorithm –NSGAI. The basic model presented in the current research is the mean-variance cross-sectional performance model of fuzzy Markowitz data envelopment analysis.

Suppose $\tilde{A} = (a, b, \alpha, \beta)$ is a trapezoidal fuzzy number. The cross-sectional performance model of fuzzy Markowitz MV is expressed as:

$$\begin{aligned} \text{Max } E \left(\sum_{i=1}^N \tilde{R}_i w_i \right) &= \sum_{i=1}^N \frac{1}{2} \left[a_i + b_i + \frac{1}{3} (\beta_i - \alpha_i) \right] w_i, \\ \text{Min } \sigma^2 \left(\sum_{i=1}^N \tilde{R}_i w_i \right) &= \left(\sum_{i=1}^N \frac{1}{2} [b_i - a_i + \frac{1}{3} (\alpha_i + \beta_i)] w_i \right)^2 + \frac{1}{72} \left[\sum_{i=1}^N (\alpha_i + \beta_i) w_i \right]^2, \end{aligned}$$

$$\begin{aligned} & \text{Max } \sum_{i=1}^N w_i \bar{e}_i, \\ & \text{s. t:} \end{aligned}$$

$$\sum_{i=1}^N z_i \leq h,$$

$$l_i z_i \leq w_i \leq u_i z_i \quad i = 1, \dots, N,$$

$$\sum_{i=1}^N w_i = 1,$$

$$w_i \geq 0 \quad i = 1, \dots, N.$$

Chen et al. [16] illustrated a comprehensive model for selecting a multi-objective portfolio in a fuzzy environment by combining the semi-variance mean model and the cross-sectional data envelopment analysis model. Then, the proposed model was re-formulated with the Sharp ratio by considering resource constraints as well as other constraints. The sharp factor model is expressed as single-factor and multi-factor models. In the single-factor model, it is assumed that the returns of all securities are correlated with each other for only one reason as a common factor to which all securities react with varying degrees of intensity. This common factor is usually considered as the market basket. Due to the fact that in the return study, the positive deviation from the average as a profit is more than expected and this issue is considered as a positive criterion, so to achieve a more consistent result with reality, semi-variance models are used, where only the negative deviation from the expected return is minimized. These models are called the mean-half variance model of portfolio optimization.

Besides the half-variance deviation, there are some downside risks. To assess the Portfolio Performance (PE), Chen and Guy [15] considered three types of data envelopment analysis approaches based on fuzzy portfolio evaluation models and based on the size of different risk scales, namely the probable variance, probable semivariance, and probable semi-absolute deviation.

6. Conclusion and Suggestion

In today's world, due to the variety of choices in each field, reviewing and selecting the best options in each field is of particular importance and this issue is much more significant for everyone to choose the best investment portfolio. To select the best investment portfolio in the general sense and the best stock portfolio, in particular, there are various criteria and indicators. Concentrating on the financial structure of companies, sales, profitability, business environment, various business, political risks, etc. can be effective in choosing a portfolio. Nowadays, according to the development of mathematical models, the use of conceptual mathematical models helps investors to be better informed about the returns and risk of different stock portfolios based on the patterns and principles of mathematical models and let them choose the best type of investment according to their standards as well as their investment policies. Due to the unique features of data envelopment analysis, it has attracted the interest of many researchers among a wide variety of mathematical models to study and select the best portfolios. Despite the ambiguous nature of the data,

combining this analysis and Fuzzy concepts leads to coming closer to reality in the models such as the Markowitz model. This paper provides an overview of some of the research conducted in optimizing investment portfolios using data envelopment analysis models, Markowitz model and the concept of fuzzy data. With the advancement of fuzzy concepts and finding new definitions of fuzzy concepts that bring issues and models closer to the realities of society, the presented models can be transformed into a variety of new fuzzy concepts and further studied. One of the new definitions and developments of the fuzzy concept is the description and concept of intuitive fuzzy in which the non-membership function is considered in addition to the membership function in the fuzzy concept and interfered in the relevant calculations. Given that based on field research, very little research has been done in the field of portfolio optimization using the Markowitz model and intuitive fuzzy concepts, more appropriate choices can be suggested to investors for future research by developing portfolio fuzzy optimization models to models with intuitive fuzzy data.

References

- [1] Heybati, F., & Haddadzadeh, R. (2008). Portfolio optimization using markowitz's mean-semi variance method on tehran stock exchange. *Future studies management*, 1(19), 39-55. (In Persian)
- [2] Rotela Junior, P., Rocha, L. C. S., Aquila, G., Balestrassi, P. P., Peruchi, R. S., & Lacerda, L. S. (2017). Entropic data envelopment analysis: a diversification approach for portfolio optimization. *Entropy*, 19(9), 352.
- [3] Tavana, M., Keramatpour, M., Santos-Arteaga, F. J., & Ghorbaniane, E. (2015). A fuzzy hybrid project portfolio selection method using data envelopment analysis, TOPSIS and integer programming. *Expert systems with applications*, 42(22), 8432-8444.
- [4] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078-1092.
- [5] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338-353.
- [6] Atanassov, K. (2016). Intuitionistic fuzzy sets. *International Journal Bioautomation*, 20, 1.
- [7] Edalatpanah, S. A. (2019). A data envelopment analysis model with triangular intuitionistic fuzzy numbers. *International journal of data envelopment analysis*, 7(4), 47-58.
- [8] Markowitz, H. M. (1978). *Portfolio selection*. Wiley, New York.
- [9] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6), 429-444.
- [10] Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the royal statistical society: series A (General)*, 120(3), 253-281.
- [11] Edalatpanah, S. A. (2018). Neutrosophic perspective on DEA. *Journal of applied research on industrial engineering*, 5(4), 339-345.
- [12] Sharifghazvini, M. R., Ghezavati, V. R., Raissi, S., & Makui, A. (2018). Integration of a new mcdm approach based on the dea, fanp with monlp for efficiency-risk assessment to optimize project portfolio by branch and bound: a real case-study. *Economic computation & economic cybernetics studies & research*, 52(1).
- [13] Ahmadi, M., Osman, M. H. M., & Aghdam, M. M. (2020). Integrated exploratory factor analysis and Data Envelopment Analysis to evaluate balanced ambidexterity fostering innovation in manufacturing SMEs. *Asia pacific management review*, 25(3), 142-155.
- [14] Mashayekhi, Z., & Omrani, H. (2016). An integrated multi-objective Markowitz-DEA cross-efficiency model with fuzzy returns for portfolio selection problem. *Applied soft computing*, 38, 1-9.
- [15] Chen, W., Gai, Y., & Gupta, P. (2018). Efficiency evaluation of fuzzy portfolio in different risk measures via DEA. *Annals of operations research*, 269(1-2), 103-127.
- [16] Chen, W., Li, S. S., Zhang, J., & Mehlatat, M. K. (2020). A comprehensive model for fuzzy multi-objective portfolio selection based on DEA cross-efficiency model. *Soft computing*, 24(4), 2515-2526.