Stock Market Portfolio Selection by Linear Programming

Conference Paper · November 2020

Hesham Alfares
King Fahd University of Petroleum and Minerals

DECISION SCIENCES INSTITUTE
Stock Market Portfolio Selection Using Linear Programming

Hesham K. Alfares
King Fahd University of Petroleum & Minerals
Email: alfares@kfupm.edu.sa

Mohammed A. Al-Marhoun
King Fahd University of Petroleum & Minerals
Email: mohammed.almarhun@gmail.com

ABSTRACT

This paper presents an integer linear programming model to optimize portfolio investment in the stock market for maximizing expected return. The model considers risk tolerance, expected dividends, and limits on both the capital investment and the number of stocks for each company. As a new feature, the model ensures that the companies whose shares are selected in the portfolio produce a sufficient variety of different products. A case study is presented, illustrating the model's application to optimize investment in shares of petrochemical companies in Saudi Arabia. Sensitivity analysis is performed to characterize the model's performance, and to make appropriate conclusions.

KEYWORDS: Portfolio selection, Linear programming, Investment Analysis, Risk Management

INTRODUCTION

Strategic investment decisions are difficult, especially in the stock market shares. This is due to the high volatility and uncertainty of the stock market, and its sensitivity to various unpredictable economic, political, and technological factors. The degree of difficulty of such decisions intensifies with larger capital investments and longer-term investment horizons. Several uncertain objectives must be balanced, including expected return and expected risk. In order to deal with risk, it is necessary to diversify the investment. The usual diversification strategy is to invest in several companies by purchasing a portfolio (group) of several company stocks. This paper adds a new dimension of diversification, by requiring the investment in several products. In other words, the companies included in the portfolio must produce a sufficient variety of different products. The proposed diversification approach thus avoids both dependence on a single company and dependence on a single product.

The portfolio selection is an important investment problem for both individuals and organizations such as banks and brokerage firms. Given a limited amount of capital, the main objective is to divide capital among different investment options to maximize the expected return resulting from the resale of purchased assets. Other investment objectives include high dividends, low risk, and fast growth. Different investors may emphasize different objectives according to their financial situations, risk-taking preferences, and investment time horizons. Since high return is associated with high risk, the fundamental question in the portfolio selection problem is to find the optimum balance between risk and return.

The classical portfolio problem either maximizes return subject to a given risk limit or minimizes risk subject to a given return limit. A popular approach to reduce investment risk is to diversify by buying the stocks of several companies. Sometimes, however, this approach may not work if all of the companies in the portfolio produce one product or a few similar products. This is applicable if the investor limits the portfolio selection to a specific market sector, such as construction, transportation, or IT. In that case, a downturn in the demand or prices of the limited set of products will negatively affect the stock prices of all companies in the portfolio. To overcome this limitation, this paper proposes two-fold diversification by investing in multiple products as well as multiple companies.

This paper presents a novel mixed-integer linear programming of the portfolio selection problem with two diversification dimensions: number of companies, and number of products. As far as the authors know, this is a new approach that has not been addressed previously in published literature. The model integrates several practical factors, including risk allowance, dividends income, and bounds on both the capital investment and the number of stocks for any individual company. To demonstrate the model's applicability and usefulness, it is applied to a real case of investment in the petrochemical sector companies in the Saudi Arabian stock market.
The remainder of this paper is organized as follows. Relevant recent literature is reviewed in Section 2. The integer programming model is formulated in Section 3. The case study application is presented in Section 4. Finally, conclusions and recommendations for future research are provided in Section 5.

LITERATURE REVIEW

Portfolio selection problem is a classical and well-studied problem in economics, finance, and operations research problem. Markowitz (1952) proposed the original mean-variance portfolio model, in which he mathematically proved how diversification can reduce the investment risk. Defining risk as the standard deviation of returns, he expressed the variance of the whole portfolio's return as a function of the individual variances of the returns of selected stocks. Using Markowitz portfolio selection model, investors can either maximize return for a given risk, or minimize risk for given return. By varying the levels of given risks and returns, the Pareto efficient front of investment alternatives can be identified. Depending on their personal levels of risk-taking or risk-aversion, each investor can a select an appropriate point on the efficient front corresponding to a specific investment portfolio (Alvani et al., 2004).

Over the years, there have been numerous extensions, variations, and applications of the Markowitz original model. In different occasions, these works have been reviewed and classified in several comprehensive surveys of literature, such those by Markowitz (1999), Rubinstein (2002), and Zhang et al. (2018). Zhang et al. (2018) classify and review several types of improvements of Markowitz's mean–variance model. These improved models include dynamic, robust, and fuzzy portfolio optimization, in addition to portfolio optimization with practical factors. The focus of this section is recent linear programming and diversification approaches to the portfolio selection problem, especially in the stock market investment decisions.

In the literature, portfolio diversification to reduce investment risk is usually considered only across firms and across geographical regions (Luigi and Jappelli, 2008). Diversification across firms means buying assets of several companies, while geographical (international) diversification means buying assets in several countries. Hui (2005) uses the statistical technique of factor analysis to select international markets to invest in. For investors interested in international diversification, he analyzes co-movement and interdependence between the US and the Asia-Pacific stock markets. Topaloglou et al. (2008) develop a stochastic dynamic programming model to optimize an international portfolio selection problem. Considering multiple periods with uncertain asset prices and currency exchange rates, the model divides available capital among different markets, and selects the assets to buy in each market.

In addition to the number of companies included in the portfolio, entropy has been used as an alternative measure of diversification. Bera and Park (2008) maximize cross-entropy as the objective function, subject to constraints derived from the mean–variance matrix. Huang (2012) develops an entropy-based method to maximize diversification in a fuzzy portfolio selection problem, using triangular fuzzy numbers to replace variances by semi-variances. Yu et al. (2014) use multiple criteria to compare the dynamic performance of several entropy-based models in the portfolio selection problem. The comparison shows that models with entropy have better diversity and asset allocation than models without entropy.

Linear programming (LP) and other optimization approaches have been extensively used to solve various versions of the portfolio selection problem. Mansini et al. (2014) survey and classify LP-based portfolio optimization techniques since 1994. Many of these are mixed integer linear programming (MILP) techniques, used after linearizing Markowitz quadratic programming model of the portfolio selection problem. Sawik (2013) reviews linear and mixed integer programming techniques for portfolio selection optimization with multiple objectives. Multi-criteria portfolio optimization techniques are classified into weighting methods, lexicographic methods, and reference point methods. Li et al. (2000) formulate a portfolio selection with transaction costs as a parametric non-quadratic programming model. Using a linear approximation of the utility function, the model is transformed to a quadratic programming model and then solved by linear programming.

Papahristodoulou and Dotzauer (2004) use LP to compare three formulations of the portfolio selection problem: (i) classical quadratic programming model (ii) maximin model, and (iii) minimum mean absolute deviation model. Approximating the variance by the conditional value at risk, Mansini et al. (2007) formulate an LP model of the portfolio optimization problem and analyze its risk aversion performance. Sadati and Nematian (2013) use an LP model to represent a fuzzy portfolio selection problem aiming to maximize the degree of both the possibility and the necessity. The model is optimally solved by using a two-level LP procedure that calculate two bounds on the objective function value. Xidonas et al. (2018) integrate three objectives and several real-world constraints in a portfolio LP model, which they apply to the European stock market.

Based on the above literature review, the model presented in this paper has a unique feature, which is two-dimensional diversification: both across firms and across products. The problem is defined and the model is formulated in the next section.

PROBLEM DEFINITION AND FORMULATION

The problem considered here is to optimize portfolio selection for a single investor with a limited budget and many investment options to choose from. These options vary widely in terms of expected returns (resale profits), expected risks, and expected dividends. The objective is to maximize total expected return subject to risk and budget limits, in addition to several diversification conditions. The problem is formulated as a linear programming model, which is presented below.

Indices

i Company index (i = 1, 2,…, *I*), where *I* is the total number of companies.
j Product index (j = 1, 2,….., *J*), where *J* is the total number of products.

Input Parameters

R_i Return of buying one share in company i.
C_i Cost of buying one share in company i, excluding commission.
D_i Dividend of one share in company i.
S_i Risk of buying one share in company i.
G_j Set of indices *i* of companies that make product j.
b Trading Commission rate in Saudi Stock Market.
B Total Budget.
L Minimum number of products in the portfolio.
d Investor minimum dividend income, as a percentage of the investment.
s Investor maximum expected risk per share.
p Maximum cost proportion of each company out of the total cost of the portfolio.
M A large positive number.
N Minimum number of shares to buy from each company selected in the portfolio.

Decision Variables

TC Total cost of the investment, excluding purchase commission.
TR Total return of the investment.
X_i Number of company i shares to buy
Y_i Equal to 1 if shares of company *i* are bought; equal to 0 otherwise.
Z_j Equal to 1 if shares of companies producing product *j* are bought, equal to 0 otherwise.

Objective Function

The objective function (1) of the portfolio selection model is to maximize *TR*, the total return of the whole investment.

$$Maximize\ TR = \sum_{i=1}^{i} R_i X_i \tag{1}$$

Constraints

Objective function (1) is optimized subject to constraints (2)-(13) listed below. Constraint (2) defines the total cost *TC* as the sum of purchase costs of all company shares. Constraint (3) ensures the sum of purchase and commission costs does not exceed the available budget *B*. Constraint (4) guarantees the total income from dividends is at least equal to a given proportion *d* of the purchase cost. Constraint (5) assures the average risk index per share does not exceed a given threshold *s*. Constraints (6) make sure the investment in each individual company not exceed a given proportion *p* of the total investment.

The two dimensions of portfolio diversification are imposed by constraints (7)-(11). Constraints (7) relate variables X_i and Y_i, by making sure that Y_i is equal to 1 if $X_i > 0$, i.e. if shares are bought of company *i*. Constraints (8) impose a lower bound, *N*, on the number of shares bought from each selected company. Constraints (9) and (10) are logical constraints relating variables Y_i and Z_j. The set of constraints (9) and (10) ensures that Z_j is equal to 1 if shares are bought of any company producing product *j*, and that Z_j is equal to 0 if no such shares are bought. Constraint (11) impose a lower bound, *L*, on the number of the number of products produced by the companies selected in the portfolio. Finally, constraint (12) specifies non-negativity and binary restrictions on relevant decision variables.

$$TC = \sum_{i=1}^{i} C_i X_i \tag{2}$$

$$TC\,(1+b) \le B \tag{3}$$

$$\sum_{i=1}^{i} D_i X_i \ge d\,TC \tag{4}$$

$$\sum_{i=1}^{i} S_i X_i \le s \sum_{i=1}^{i} X_i \tag{5}$$

$$C_i X_i \le p\,TC \qquad for\ (i = 1,2,\dots\dots,I) \tag{6}$$

$$X_i \le M\,Y_i \qquad for\ (i = 1,2,\dots\dots,I) \tag{7}$$

$$X_i \ge N\,Y_i \qquad for\ (i = 1,2,\dots\dots,I) \tag{8}$$

$$\sum_{i \in G_j}^{j} Y_i \ge Z_j, \qquad for\ (j = 1,\dots,J) \tag{9}$$

$$\sum_{i \in G_j}^{j} Y_i \le MZ_j, \qquad for\ (j = 1,\dots,J) \tag{10}$$

$$\sum_{j=1}^{J} Z_j \ge L \tag{11}$$

$$TC \ge 0, \quad X_i \ge 0, \quad Y_i = (0 \text{ or } 1), \quad Z_j = (0 \text{ or } 1)$$

$$for\ (i = 1,2,\dots\dots,I), \qquad (j = 1,2,\dots\dots,J) \tag{12}$$

CASE STUDY

The above integer programming model has been applied for portfolio selection of investment in the stock market in Saudi Arabia. In particular, the case study application is limited to 10 publicly-owned petrochemical companies whose shares are traded in the Saudi Arabian stock market. For investors interested in a single market sector, investing in several companies may not provide sufficient diversification. This is because the selected companies in the same economic sector may produce a limited number of similar products that face the same set of market forces. Therefore, it is important to have a second dimension of diversification, namely diversification in the number of products made by the selected companies. Using real historical data, the inputs for the model are prepared and presented below.

First, relevant publicly available data from Bloomberg and the Saudi Exchange (Tawaul) Websites was collected for the last three years (2016-2018) for the 10 companies under study. Unfortunately, this is the only period during which all the data required for the model can be extracted. From this data, the expected returns R_i and expected dividends D_i were calculated as the mean returns and dividends during these three years. For the unit stock prices C_i, we simply used the latest values available at the end of 2018. To determine the expected risk for each company, S_i, input positive differences (expected return - actual return) for each year were averaged for the last three years, and then the average market risk for the petrochemical sector was added. The annual values of the expected returns, as well as the average market risk value, were provided by expert financial analysts. All of these input data values are shown in Table 1.

Another important component of input data is the set of products produced by the 10 companies included in the case study. Altogether, these 10 companies produce a total of 22 petrochemical products that are listed in Table 2. The company-product matrix in shown in Table 3. The rows in Table 3 show the set of companies that make each product, while columns show the set of products made by each company. It should be noted that the number “1” in a cell in Table 3 indicates the given (row) company makes the corresponding (column) product, while a “0” in a cell indicates the company does not. Therefore, for each column, the set of indices *i* of rows with “1” represents the set G_j of all companies that make a given product *j*.

The final input parameter relevant to the Saudi stock market is b, the trading commission rate, which is 0.16% of the stock purchase cost. The five remaining input parameters are flexible, because they represent the values to be specified by each individual investor. These are the values of parameters *B*, *L*, *d*, *s*, and *p*. For the case study, the specific values of these parameters are given in Table 4. These values are typical for a class of high-end investors in the Saudi stock market.

Solving the linear programming model of the portfolio selection problem with the values shown in Table 4 leads to the optimum solution shown in Table 5.

For the optimum number of shares shown in Table 5, the corresponding objective function, i.e. the expected return *TR*, is equal to 142,669.87. The optimum solution calls for selecting four out of the ten petrochemical companies to invest in. Although shares are bought in only four companies, the total number of products made by these companies is 14. For this optimum investment, the income from dividends is equal to 1.13% of the total cost *TC*, which is greater than the minimum proportion, *d* = 1%. Moreover, the average risk index per share is equal to 7.51%, which is less than given threshold, *s* = 10%.

Table 1. Petrochemical sector companies and their input data values

i	Company Name	Abbreviation	C_i	D_i	R_i	S_i
1	National Industrialization Co	TASNEE	18.5	0.00	0.29	0.085
2	Saudi Arabian Fertilizer Co	SAFCO	73.5	2.25	0.15	0.051
3	Yanbu National Petrochemicals Co	YANSAB	62.5	3.33	0.15	0.101
4	Advanced Petrochemical Co	APC	52.75	2.56	0.14	0.070
5	Saudi Basic Industries Corp	SABIC	98.5	4.20	0.05	0.060
6	Saudi Arabian Mining Co	MAADEN	48.5	0.00	0.04	0.067
7	Saudi International Petrochemical Co	SIPCHEM	19.5	0.55	0.29	0.040
8	National Petrochemical Co	PETROCHEM	26	0.50	0.25	0.085
9	Nama Chemicals Co	NAMA	20.25	0.58	0.20	0.106
10	Saudi Kayan Petrochemical Co	KAYAN	14.5	0.00	0.20	0.086

Table 2. list of products produced by the petrochemical companies.

j	Product	*j*	Product
1	Polyethylene	12	Polycarbonate
2	Polypropylene	13	Bisphenol A
3	Polystyrene	14	Diammonium Phosphate (DAP)
4	Mono Ethylene Glycol (MEG)	15	Epoxy Resin
5	Methyl Tert-Butyl Ether (MTBE)	16	Hydrochloric Acid
6	Benzene	27	Liquid Caustic Soda
7	Urea	28	Soda Granule
8	Ammonia	29	Methanol
9	Polyvinyl Chloride (PVC)	20	Butanol
10	Purified Terephthalic Acid (PTA)	21	Acetic Acid
11	Propylene	22	Vinyl Acetate Monomer (VAM)

Table 3. Matrix showing the relationship between each company (*i*) and each product (*j*)

i \ *j*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
7	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
8	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
10	1	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0

Table 4. Given input values for an individual investor

s	*d*	*p*	*L*	*B*
10%	1%	30%	12	10 Million

Table 5. Optimum number of shares to buy for each company

i	1	2	3	4	5	6	7	8	9	10
X_i	161,903	0	0	0	0	0	153,600	0	49,305	206,565

CONCLUSIONS

This paper considered the portfolio selection problem with two dimensions of diversification: diversification across companies, and diversification across products. A linear programming (LP) model has been formulated to maximize expected return. In addition to enforcing two dimensions of diversification, several practical constraints are incorporated in the LP model, including budget, dividends incomes, risk index, and limits on individual company's number of shares and proportion of the total investment. The model has been successfully applied to a case study involving investment in petrochemical companies listed in the Saudi Stock Market. For future research, the LP model can be extended to consider multiple time periods, multiple objectives, or multiple international markets. Possibly, new dimensions of portfolio diversification may also be proposed.

REFERENCES

Alvani, S. M., Azar, A., & Danayeefar, H. (2004). *Quantitative Research Methodology in Management: A Comprehensive Approach*. Saffar, Tehran, Iran.

Bera, A. K., & Park, S. Y. (2008). Optimal portfolio diversification using the maximum entropy principle. *Econometric Reviews*, 27(4-6), 484-512.

Bloomberg professional services. https://www.bloomberg.com/professional/solution/content-and-data/.

Huang, X. (2012). An entropy method for diversified fuzzy portfolio selection. *International Journal of Fuzzy Systems*, 14(1), 160-165.

Hui, T. K. (2005). Portfolio diversification: a factor analysis approach. *Applied Financial Economics*, 15(12), 821-834.

Li, Z. F., Wang, S. Y., & Deng, X. T. (2000). A linear programming algorithm for optimal portfolio selection with transaction costs. *International Journal of Systems Science*, 31(1), 107-117.

Luigi, G., & Jappelli, T. (2008). *Financial Literacy and Portfolio Diversification*. ECO 2008/31, European University Institute (EUI), Fiesole, Italy.

Mansini, R., Ogryczak, W., & Speranza, M. G. (2007). Conditional value at risk and related linear programming models for portfolio optimization. *Annals of Operations Research*, 152(1), 227-256.

Mansini, R., Ogryczak, W., & Speranza, M. G. (2014). Twenty years of linear programming based portfolio optimization. *European Journal of Operational Research*, 234(2), 518-535.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77-91.

Markowitz, H. M. (1999). The early history of portfolio theory: 1600–1960. *Financial Analysts Journal*, 55(4), 5-16.

Papahristodoulou, C., & Dotzauer, E. (2004). Optimal portfolios using linear programming models. *Journal of the Operational Research Society*, 55(11), 1169-1177.

Rubinstein, M. (2002). Markowitz's “Portfolio Selection”: A Fifty-Year Retrospective. *The Journal of Finance*, 57(3), 1041-1045.

Sadati, M. E. H., & Nematian, J. (2013). Two-level linear programming for fuzzy random portfolio optimization through possibility and necessity-based model. *Procedia Economics and Finance*, 5, 657-666.

Sawik, B. (2013). Survey of multi-objective portfolio optimization by linear and mixed integer programming. In *Applications of Management Science* (pp. 55-79). Emerald Group Publishing Limited.

Tadawul website. https://www.tadawul.com.sa/wps/portal/tadawul/home/.

Topaloglou, N., Vladimirou, H., & Zenios, S. A. (2008). A dynamic stochastic programming model for international portfolio management. *European Journal of Operational Research*, 185(3), 1501-1524.

Xidonas, P., Hassapis, C., Mavrotas, G., Staikouras, C., & Zopounidis, C. (2018). Multiobjective portfolio optimization: bridging mathematical theory with asset management practice. *Annals of Operations Research*, 267(1-2), 585-606.

Yu, J. R., Lee, W. Y., & Chiou, W. J. P. (2014). Diversified portfolios with different entropy measures. *Applied Mathematics and Computation*, 241, 47-63.

Zhang, Y., Li, X., & Guo, S. (2018). Portfolio selection problems with Markowitz's mean–variance framework: a review of literature. *Fuzzy Optimization and Decision Making*, 17(2), 125-158.