

# An Alternative Proof to Markowitz's Model

Yaniv Zaks\*

Department of Mathematics, Bar-Ilan University, Israel

\*Corresponding author: yaniv.zaks@gmail.com

Received May 07, 2013; Revised June 20, 2013; Accepted June 21, 2013

**Abstract** In the fundamental paper on portfolio selection, Markowitz (1952) described via geometric reasoning his innovative theory and provided the explicit optimal selection for the cases of 3 and 4 assets. Merton (1972) obtained for the general case the efficient portfolio frontiers explicitly by using Lagrange multipliers. In this paper, we suggest a geometric approach to achieve the explicit optimal selection for the general case thus generalizing Markowitz's original approach to achieve the explicit presentation of the desired selection.

**Keywords:** portfolio selection, Markowitz model, quadratic programming, euclidean projection

## 1. Introduction

The basis of the portfolio selection problem was first defined by the Nobel laureate Harry Markowitz in 1952. In his famous paper he combined, for the first time, the mean and the variance of the return in order to select an optimal portfolio. The variance stands for measuring the risk of the portfolio. The Nobel laureate Merton H. Miller referred in his essay "The history of finance: an eyewitness account" to Markowitz's model and wrote "Markowitz in this remarkable paper gave, for the first time, a precise definition of what had hitherto been just vague buzzwords: risk and return". Thus, it is not surprising that Markowitz's model is the first model that every student learns in the field of portfolio selection. Many times this is also the only model that is studied.

In this paper, we give an alternative proof to Markowitz's model. Markowitz (1952) explained his innovative theory by geometric reasoning and provided the explicit optimal selection for the cases of 3 and 4 assets. Twenty years later, Merton (1972) used the Lagrange multipliers method to obtain an explicit formula for the efficient portfolio frontier for the general case. Solving the Lagrange first order condition gives us a critical point that might be the desired minimum, but theoretically it might be a maximum or a saddle point as well. While teaching this model in the classroom, there is usually no discussion as to the character of the critical point. In order to verify that we have obtained a minimum, we should verify that some extra conditions hold. Unfortunately, this kind of discussion is very rare in the classroom due to its complexity.

We suggest a geometric approach to achieve the explicit optimal portfolio for the general case; thus we follow Markowitz's original presentation of his theory. Following the geometric approach, it is easy to explain that there exist a solution and to show that this solution is unique. To derive the optimal solution we consider well-known formulas in analytic geometry.

It is important to mention that in our analysis it is essential to assume that the covariance matrix is a positive definite matrix. This assumption is required by Merton as well. The two techniques, obviously, lead to the same single solution. According to the method we are using to solve the model in this paper, there is no need for extra calculations or explanations in order to verify that the obtained solution is the minimum. This is the major advantage of our approach over the Lagrange multipliers method.

We will present the objective function of Markowitz's model in terms of a Euclidean projection of a point on a hyperplane. A hyperplane of an  $n$ -dimensional space is a subspace with dimension  $n-1$ , for example, a line in a plane ( $\mathbb{R}^2$ ) or a plane in  $\mathbb{R}^3$ . A hyperplane is determined by a vector  $\mathbf{a} \in \mathbb{R}^n$  and a scalar  $b$ . In mathematical notations, a hyperplane is defined as the following set

$$H_{a,b} = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{a} = b \}$$

where  $\mathbf{x}^T \mathbf{a}$  is the inner product of  $\mathbf{x}$  and  $\mathbf{a}$ , i.e.  $\mathbf{x}^T \mathbf{a} = \langle \mathbf{x}, \mathbf{a} \rangle = \sum_{i=1}^n x_i a_i$ . Note that for any two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we use the notations  $\mathbf{x}^T \mathbf{a}$  and  $\langle \mathbf{x}, \mathbf{a} \rangle$  alternately throughout the rest of the paper.

Let  $\mathbf{x}, \mathbf{c} \in \mathbb{R}^n$ . The Euclidean distance between  $\mathbf{x}$  and  $\mathbf{c}$  is denoted by  $\| \mathbf{x} - \mathbf{c} \|$  such that

$$\| \mathbf{x} - \mathbf{c} \|^2 = \langle \mathbf{x} - \mathbf{c}, \mathbf{x} - \mathbf{c} \rangle = \sum_{i=1}^n (x_i - c_i)^2$$

A Euclidean projection of a point  $\mathbf{c} \in \mathbb{R}^n$  on the hyperplane  $H_{a,b}$  is a point that belongs to  $H_{a,b}$  that is closest to the point  $\mathbf{c}$  (in the sense of Euclidean distance). It is well known in analytic geometry that such a point exists and it is unique. Moreover, there is an explicit solution for the projection. In mathematical notations, we state the problem of finding the Euclidean projection as the following optimization problem:

$$\begin{cases} \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{x} - \mathbf{c}\|^2 \\ \text{s. t. } \mathbf{x}^T \mathbf{a} = b \end{cases}$$

and the optimal solution, i.e., the projection, is:

$$\mathbf{x}^* = \frac{b - \mathbf{c}^T \mathbf{a}}{\mathbf{a}^T \mathbf{a}} \mathbf{a} + \mathbf{c} \quad (1)$$

## 2. The Optimal Solution for Markowitz Model

The Markowitz model takes into consideration the expected value  $\mu_x$  and the variance  $\sigma_x^2$  of a portfolio  $R_x = \mathbf{x}^T \mathbf{R}$ , where  $\mathbf{R}$  is the vector of returns and  $\mathbf{x}$  is the weights vector of the portfolio. The optimization problem in the Markowitz model is the following:

$$\begin{cases} \max_{\mathbf{x} \in \mathbb{R}^n} \{2\tau\mu_x - \sigma_x^2\} \\ \text{s. t. } \mathbf{1}^T \mathbf{x} = 1 \end{cases} \quad (2)$$

where  $\tau$  is the risk tolerance of the investor.

It is well known (Panjer et al. (2001), Best & Grauer (1992)) that the optimal solution is the following:

$$\mathbf{x}^* = \frac{1}{\mathbf{1}^T \sum_R^{-1} \mathbf{1}} \sum_R^{-1} \mathbf{1} + \tau \left( \sum_R^{-1} \boldsymbol{\mu}_R - \frac{\mathbf{1}^T \sum_R^{-1} \boldsymbol{\mu}_R}{\mathbf{1}^T \sum_R^{-1} \mathbf{1}} \sum_R^{-1} \mathbf{1} \right) \quad (3)$$

Where  $\boldsymbol{\mu}_R = (E[R_1], \dots, E[R_n])^T$  is the vector of expectations and  $\sum_R$  is the covariance matrix of the returns. Furthermore,  $\mu_x = \mathbf{x}^T \boldsymbol{\mu}_R$  and  $\sigma_x^2 = \mathbf{x}^T \sum_R \mathbf{x}$ . By changing the parameter  $\tau$ , we obtain the efficient frontier. In our analysis we refer to the equivalent minimization problem:

$$\begin{cases} \min_{\mathbf{x} \in \mathbb{R}^n} \{\sigma_x^2 - 2\tau\mu_x\} \\ \text{s. t. } \mathbf{1}^T \mathbf{x} = 1 \end{cases} \quad (4)$$

The covariance matrix  $\sum_R$  is a symmetric matrix and we assume that it is a positive definite matrix as well. Hence there exists a matrix  $\mathbf{G}$  such that  $\sum_R = \mathbf{G}^2$ ,  $\mathbf{G} = \mathbf{G}^T$  and  $\mathbf{G}^{-1}$  exists. It follows that

$$\begin{aligned} & \left\| \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 \\ &= \left\langle \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\rangle \\ &= \langle \mathbf{G}\mathbf{x}, \mathbf{G}\mathbf{x} \rangle - 2\tau \langle \mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}\mathbf{x} \rangle + \langle \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \rangle \end{aligned}$$

Hence,

$$\left\| \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 - \langle \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \rangle = \langle \mathbf{G}\mathbf{x}, \mathbf{G}\mathbf{x} \rangle - 2\tau \langle \mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}\mathbf{x} \rangle \quad (5)$$

By the equality  $\mathbf{G} = \mathbf{G}^T$ , it follows that the first term on the right hand side satisfies:

$$\langle \mathbf{G}\mathbf{x}, \mathbf{G}\mathbf{x} \rangle = \mathbf{x}^T \mathbf{G}^T \mathbf{G}\mathbf{x} = \mathbf{x}^T \sum_R \mathbf{x} \quad (6)$$

and the second term satisfies:

$$2\tau \langle \mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}\mathbf{x} \rangle = 2\tau \mathbf{x}^T \mathbf{G}^T \mathbf{G}^{-1}\boldsymbol{\mu}_R = 2\tau \mathbf{x}^T \boldsymbol{\mu}_R = 2\tau\mu_x \quad (7)$$

Hence, the following equality holds:

$$\sigma_x^2 - 2\tau\mu_x = \left\| \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 - \langle \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \rangle$$

where the last term on the right hand side does not depend on  $\mathbf{x}$ . Therefore, the optimization problem

$\min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \left\| \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 \right\}$  has the same optimal solution as  $\min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \sigma_x^2 - 2\tau\mu_x \right\}$ .

It remains to solve the following optimization problem:

$$\begin{cases} \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \left\| \mathbf{G}\mathbf{x} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 \right\} \\ \text{s. t. } \mathbf{1}^T \mathbf{x} = 1 \end{cases} \quad (8)$$

Let  $\mathbf{y} = \mathbf{G}\mathbf{x}$ . Substituting  $\mathbf{G}\mathbf{x}$  by  $\mathbf{y}$  in(8) leads to:

$$\begin{cases} \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \left\| \mathbf{y} - \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \right\|^2 \right\} \\ \text{s. t. } \mathbf{1}^T \mathbf{G}^{-1}\mathbf{y} = 1 \end{cases} \quad (9)$$

By the symmetry of  $\mathbf{G}$ , it is possible to write the constraint as  $\langle \mathbf{G}^{-1}\mathbf{1}, \mathbf{y} \rangle = 1$ . We obtain that the optimization problem in (9) is equivalent to the problem of finding the Euclidean projection of the vector  $\tau\mathbf{G}^{-1}\boldsymbol{\mu}_R$  onto the hyperplane  $\langle \mathbf{G}^{-1}\mathbf{1}, \mathbf{y} \rangle = 1$ . This problem has the unique solution:

$$\mathbf{y} = \frac{1 - \langle \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}^{-1}\mathbf{1} \rangle}{\langle \mathbf{G}^{-1}\mathbf{1}, \mathbf{G}^{-1}\mathbf{1} \rangle} \mathbf{G}^{-1}\mathbf{1} + \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R \quad (10)$$

Notice that  $\langle \tau\mathbf{G}^{-1}\boldsymbol{\mu}_R, \mathbf{G}^{-1}\mathbf{1} \rangle = \tau\boldsymbol{\mu}_R^T \sum_R^{-1} \mathbf{1} = \tau\mathbf{1}^T \sum_R^{-1} \boldsymbol{\mu}_R$  and  $\langle \mathbf{G}^{-1}\mathbf{1}, \mathbf{G}^{-1}\mathbf{1} \rangle = \mathbf{1}^T \sum_R^{-1} \mathbf{1}$ . Substituting  $\mathbf{y} = \mathbf{G}\mathbf{x}$  in(10) yields:

$$\mathbf{x} = \frac{1 - \tau\mathbf{1}^T \sum_R^{-1} \boldsymbol{\mu}_R}{\mathbf{1}^T \sum_R^{-1} \mathbf{1}} \sum_R^{-1} \mathbf{1} + \tau \sum_R^{-1} \boldsymbol{\mu}_R \quad (11)$$

as required. Note that in the above proof, we did not consider any derivatives.

## 3. Conclusions

In this paper, we solved the Markowitz's model by following a geometric approach. We showed that the original optimization problem is equivalent to the problem of finding a Euclidean projection of a point on a hyperplane. For this reason, there was no need for any further calculations in order to verify whether or not this

solution leads to the minimum of the objective function. This is the main advantage of this technique over the Lagrange multipliers method.

The proposed method may be valuable in some generalizations as well. One generalization of Markowitz model is presented in Kriens & van Lieshout (1998). In that paper, the authors replaced the expectation  $\mu_R$  and the covariance matrix  $\Sigma_R$  by the expectation and the covariance matrix of a non-linear transformation of the random vector  $R$ . They required that the new expectation and the new variance be concave and continuously differential on the set of feasible portfolios. Following the technique in this paper, we may consider any transformation that keeps the positive definite property of the covariance matrix. In addition, any transformation of the expectation is allowed. Another generalization appears in Li & Ng(2000), where the authors are seeking the optimal capital allocation in a multiperiod portfolio. Following our technique, it is easy to explain the existence of a single optimal solution.

## References

- [1] Kriens J. and van Lieshout J. Th., 1998. Notes on the Markowitz portfolio selection method. *Statistica Neerlandica* 42 (3): 181-191.
- [2] Li D., Ng W.L., 2000, Optimal dynamic portfolio selection: multiperiod mean-variance formulation, *Mathematical Finance* 10(3): 387-406.
- [3] Markowitz H., 1952. Portfolio selection. *The Journal of Finance* 7 (1): 77-91.
- [4] Merton R.C., 1972. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7 (4): 1851-72.
- [5] Best M.J. and Grauer R.R., 1992. Positively weighted minimum-variance portfolios and the structure of asset expected returns. *Journal of Financial and Quantitative Analysis* 27: 513-537.
- [6] Miller M.H., 2000, The history of finance: an eyewitness account, *Journal of Applied Corporate Finance* 13(2): 8-14.
- [7] Panjer H.H., editor: Boyle P.P...[et al.], 2001, *Financial Economics: With Application to Investments, Insurance and Pensions*. The Society of Actuaries.