An alternative proof to Markowitz’s Model

Yaniv Zaks
Department of Mathematics, Bar-Ilan University

August 14, 2012

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 programing, Euclidean projection.

JEL classification: G11, C61.

1 Introduction

In the fundamental paper in portfolio selection, Markowitz (1952) explained his innovative theory by geometric reasoning 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. Solving the first order condition gives us a critical point that might be the desired maximum, but it might be a minimum or a saddle point as well. In order to verify that we obtained a maximum, there is a need to calculate the Hessian matrix or to verify that the conditions of Kuhn-Tucker theorem hold. When this topic is studied at class, it is very common to stop the proof after deriving the solution of the first order condition.

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

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

An Euclidean projection of a point $c \in \mathbb{R}^n$ on the hyperplane $H_{a,b}$ is a point that belong to $H_{a,b}$ that is most close to the point $c$ (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 introduce this problem as follows:
\[
\begin{cases}
\min_x \{ \|x - c\|^2 \} \\
\text{s.t. } x^T a = b
\end{cases}
\]
and the optimal solution, i.e. the projection, is:
\[
x^* = \frac{b - c^T a}{a^T a} a + c
\]
Note that for any two vectors \( x, y \in \mathbb{R}^n \), we use the notations \( x^T y \) and \( \langle x, y \rangle \) alternately throughout the rest of the paper.

2 The optimal solution

Markowitz model takes into consideration the expected value \( \mu_x \) and the variance \( \sigma_x^2 \) of a portfolio \( R_x = x^T R \), where \( R \) is the vector of returns and \( x \) is the weights vector of the portfolio. The optimization problem in Markowitz model is as follows:
\[
\begin{cases}
\max_x \{ 2\tau \mu_x - \sigma_x^2 \} \\
\text{s.t. } 1^T x = 1
\end{cases}
\]
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 one:
\[
x^* = \frac{1}{1^T \Sigma_R^{-1} 1} \Sigma_R^{-1} 1 + \tau \left( \Sigma_R^{-1} \mu_R - \frac{1^T \Sigma_R^{-1} \mu_R}{1^T \Sigma_R^{-1} 1} 1 \right)
\]
where \( \mu_R = (E[R_1], ..., E[R_n])^T \) and \( \Sigma_R \) is the covariance matrix of the returns. By changing the parameter \( \tau \) we obtain the efficient frontier.

In our analysis we refer to the equivalent minimization problem:
\[
\begin{cases}
\min_x \{ \sigma_x^2 - 2\tau \mu_x \} \\
\text{s.t. } 1^T x = 1
\end{cases}
\]
The covariance matrix \( \Sigma_R \) is a symmetric matrix and we assume that it is a positive definite matrix as well. Hence there exists a matrix \( G \) such that \( \Sigma_R = G^2 \), \( G = G^T \) and \( G^{-1} \) exists. It follows that
\[
\|Gx - \tau G^{-1} \mu_R\|^2 = \langle Gx - \tau G^{-1} \mu_R, Gx - \tau G^{-1} \mu_R \rangle =
\langle Gx, Gx \rangle - 2\tau \langle G^{-1} \mu_R, Gx \rangle + \langle \tau G^{-1} \mu, \tau G^{-1} \mu \rangle.
\]
Hence
\[
\|Gx - \tau G^{-1} \mu_R\|^2 = \langle \tau G^{-1} \mu, \tau G^{-1} \mu \rangle = \langle Gx, Gx \rangle - 2\tau \langle G^{-1} \mu_R, Gx \rangle.
\]
Since \( G = G^T \) it follows that the first term on the right hand side satisfies:
\[(Gx, Gx) = x^T G^T G x = x^T \Sigma_R x = \sigma_x \] (6)

and the second term satisfies:

\[2\tau (Gx, G^{-1} \mu_R) = 2\tau x^T G^{-1} \mu_R = 2\tau x^T \mu_R = 2\tau \mu_x . \] (7)

Hence, the next equality holds:

\[\sigma_x^2 - 2\tau \mu_x = \|Gx - \tau G^{-1} \mu_R\|^2 - \langle \tau G^{-1} \mu_R, \tau G^{-1} \mu_R \rangle \]

where the last term on the right-hand side does not depend on \(x\). Therefore, the optimization problem

\[\min_x \left\{ \|Gx - \tau G^{-1} \mu_R\|^2 \right\} \]

has the same optimal solution as

\[\min_x \{ \sigma_x^2 - 2\tau \mu_x \}. \]

It remains to solve the following optimization problem:

\[\min \left\{ \|Gx - \tau G^{-1} \mu_R\|^2 \right\} \quad s.t. \quad 1^T x = 1 \] (8)

Let \(y = Gx\). Substituting \(Gx\) by \(y\) in the last optimization problem leads to:

\[\min_{y} \left\{ \|y - \tau G^{-1} \mu_R\|^2 \right\} \quad s.t. \quad 1^T G^{-1} y = 1 \] (9)

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

\[y = 1 - \frac{\langle \tau G^{-1} \mu_R, G^{-1} 1 \rangle}{\langle G^{-1} 1, G^{-1} 1 \rangle} G^{-1} 1 + \tau G^{-1} \mu_R . \] (10)

Notice that

\[\langle \tau G^{-1} \mu_R, G^{-1} 1 \rangle = \tau \mu_R^T \Sigma_R^{-1} 1 = \tau 1^T \Sigma_R^{-1} \mu \] and

\[\langle G^{-1} 1, G^{-1} 1 \rangle = 1^T \Sigma_R^{-1} 1 . \]

Substituting \(y = Gx\) in (10) yields:

\[x = \frac{1 - \tau 1^T \Sigma_R^{-1} \mu_R}{1^T \Sigma_R^{-1} 1} \Sigma_R^{-1} 1 + \tau \Sigma_R^{-1} \mu_R . \] (11)

as required.

Note that in the proof above, we did not consider any derivatives. A generalization of Markowitz model is presented in Kriens & 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 are concave and continuously differential on the set of feasible portfolios. In our proof we may consider any transformation that keeps on the positive definite property of the covariance matrix. In addition, any transformation for the expectation is allowed.

References

[1] Michael J. Best and Robert R. Grauer, 1992, Positively weighted minimum-variance portfolios and the structure of asset expected returns, Journal of financial and quantitative analysis, 27, pp 513-537.
[2] Markowitz H., 1952, Portfolio selection, The journal of finance, Vol. 7, No. 1, pp. 77-91.

[3] Merton R.C., 1972, An analytic derivation of the efficient portfolio frontier, Journal of financial and quantitative analysis, Vol. 7, No. 4, pp 1851-1872.

[4] Kriens J., van Lieshout J. Th., 1998, Notes on the Markowitz portfolio selection method, Statistica Neerlandica, Vol. 42, Issue 3, pp. 181-191.

[5] Panjer H.H. editor: Boyle P.P.. [et al.], 2001, Financial economics: with application to investments, insurance and pensions, The Society of Actuaries.