6
Portfolio Optimization-II

6.1 Introduction

We have presented Markowitz's mean variance model for portfolio optimization in the last chapter. But contrary to its theoretical reputation, this model in its original form has not found much favor with the practitioners to construct large scale portfolios. There are several theoretical and practical reasons for not using this model extensively in practice - particularly when the number of assets in the portfolio is large. This chapter aims to understand these reasons and then present some other models which have been developed to improve Markowitz's model both theoretically and computationally.

6.2 Markowitz's Model: Some Theoretical and Computational Issues

Theoretically, Markowitz's model is known to be valid if returns $r_i's$ are multivariate normally distributed and the investor is *risk averse* in the sense that he/she prefers less standard deviation of the portfolio to more. But one is not fully convinced of the validity of the standard deviation as a measure of risk. An investor is certainly unhappy to have small or negative profit, but feels happy to have larger profit. In other words, this means that the investor's perception about risk is not symmetric about the mean. There are several empirical studies, which reveal that most r_i are not normally or even symmetrically distributed. In this scenario, one possible approach seems to be to consider the skewness and kurtosis of the distribution in addition to the mean and variance and extend the Markowitz's model to generate the efficient frontier in (mean, variance, skewness, kurtosis)-space. This has been the approach of some of the models, e.g. Konno and Suzuki [79], and

Joro and Na [71], but has not found much favor because the resulting optimization problem is not easy to handle. An alternative and popular approach is to introduce certain new measures of risk which carry information about the possible portfolio losses implied by the tail of the return distribution, even in the case when the distribution is not symmetric. This takes care of those situations where the return distribution is heavily tailed. These risk measures are called *downside or safety-first risk measures* which aim to maximize the probability that the portfolio loss is below a certain acceptable level, commonly referred as the benchmark or the disaster level. Thus these risk measures are *quantile based risk measures* and are different from standard deviation or other *moment based risk measures*. Some of the most popular quantile based risk measures are value at risk (VaR) and conditional value at risk (CVaR). Since downside risk measures of individual securities cannot be easily aggregated into portfolio downside risk measures (we need the entire joint distribution of security returns), their application in practice requires computationally intensive non parametric estimation, simulation and optimization techniques.

There is another major problem associated with the classical Markowitz's model. This model gives us an optimal portfolio assuming that we have perfect information about μ_i's and σ_{ij}'s for the assets that we are considering. Therefore an important practical issue is the estimation of the μ_i's and σ_{ij}'s. A reasonable approach for estimating these data is to use time series of past returns r_{it} which represents the return of i^{th} asset from time $(t-1)$ to t, where $t = 1, 2, \ldots, T$. However, it has been observed that small changes in the time series r_{it} lead to changes in the μ_i's and σ_{ij}'s that often lead to significant changes in the optimal portfolio. This is a fundamental weakness of the Markowitz model, no matter how cleverly μ_i's and σ_{ij}'s are computed. This is because the optimal portfolio construction is very sensitive to small changes in the data. Only one small change in one μ_i may produce a totally different portfolio. In fact recent research (Chopra and Ziemba [30]) has revealed that errors in the estimation of means μ_i can be more damaging than errors in other parameters. This has motivated researchers to employ *robust optimization techniques* in Markowitz's model, e.g. Ben-Tal [9], and T*ütüncü* and Koenig [139]. A much simpler approach is to consider portfolio optimization under a minimax rule (Cai et al. [23]) and provide some flexibility by allowing μ_i to lie in some interval $a_i \leq \mu_i \leq b_i$ (Deng et al. [36]).

The mean-variance model of Markowitz, in general, results in a dense quadratic programming problem. If the number of assets in the portfolio is large then it becomes very difficult to obtain an optimal solution of such large-scale dense quadratic programming problem on a real time basis. This has motivated researchers to consider mean-absolute deviation of the portfolio as a measure of risk (Konno and Yamazaki [80]) so that the resulting optimization problem reduces to a linear programming problem. These models are called L_1-*risk models* because these are based on L_1 metric on $\mathbf{R}^n$. In this terminology, Markowitz's model can be termed as a L_2-*risk model* since it uses variance as a risk measure which is based on the notion of L_2 metric. In the same spirit, the models based on the minimax rule portfolio selection strategies can be termed as L_∞-*risk models*. It is simple to note that L_1 and L_∞ risk models will result in a linear programming formulation. This is in contrast to an L_2- risk model which results in a quadratic programming formulation. Therefore computationally, L_1 and L_∞ risk models are easier to handle in comparison to L_2 risk model. The L_1, L_2 and L_∞ are three standard moment based risk models which have been studied in the literature.

The above discussion suggests that the choice of a proper risk measure is very crucial to study portfolio selection problems. But then what would be the guiding principles for this choice? The concept of *coherent risk measures* (Artzner et al. [5]) is an important contribution in this regard. This stipulates an axiomatic approach to the study of risk measure by presenting certain desirable properties.

Some other issues associated with Markowitz's model are very natural and these have been incorporated in the original model. These are with regard to its extension for the multi-period scenario and also to incorporate transaction costs. We shall discuss some of the models in the subsequent sections which addresses the above issues.

6.3 Mean Absolute Deviation Based Portfolio Optimization: A L_1-Risk Model

We first introduce the L_1-risk measure and formulate the corresponding portfolio optimization problem with this risk measure. Let an investor has initial wealth M_0 which is to be invested in n assets a_i (i=1,2,...,n). Let r_i be the return of the asset a_i which is a random variable. Also let x_i be the amount of money to be invested in the asset a_i out of the total fund M_0.

We now define $\mu_i = E(r_i)$ and $q_i = E(|r_i - \mu_i|)$, i=1,2,...,n. Then μ_i denotes the expected return rate of the asset a_i and q_i denotes the expected absolute deviation of r_i from its mean. Then the expected return of a portfolio $(x_1, x_2, \dots, x_n)$ is given by

$$\mu = E\left(\sum_{i=1}^{n} r_i x_i\right) = \sum_{i=1}^{n} \mu_i x_i.$$

We now define L_1-risk measure or the mean absolute deviation of the portfolio $(x_1, x_2, \ldots, x_n)$.

Definition 6.3.1 **(L_1-Risk Measure of a Portfolio)** *Let $(x_1, x_2, \ldots, x_n)$ be the given portfolio. Then its L_1-risk measure or mean absolute deviation is defined as*

$$w_{L_1}(x_1, x_2, \ldots, x_n) = E\left[\left|\sum_{i=1}^{n} r_i x_i - E\left(\sum_{i=1}^{n} r_i x_i\right)\right|\right].$$

In terms of the L_1-risk measure $w_{L_1}(x_1, x_2, \ldots, x_n)$, the L_1-risk model of the portfolio optimization problem is formulated as

$$\text{Min} \qquad w_{L_1}(x_1, x_2, \ldots, x_n) = E\left[\left|\sum_{i=1}^{n} (r_i - \mu_i) x_i\right|\right]$$

subject to

$$\begin{aligned} \sum_{i=1}^{n} \mu_i x_i &\geq \alpha M_0 \\ \sum_{i=1}^{n} x_i &= M_0 \\ 0 \leq x_i &\leq u_i \qquad (i = 1, 2, \ldots, n), \end{aligned} \tag{6.1}$$

where $(\alpha > 1)$ is a parameter representing the minimum rate of return required by the investor. Also u_i is the maximum amount of the money which can be invested in the asset a_i.

In (6.1) it may be noted that by defining $w_i = x_i/M_0$, $(i = 1, 2, \ldots, n)$, the portfolio $(x_1, x_2, \ldots, x_n)$ can also be represented in terms of its weight vector $(w_1, w_2, \ldots, w_n)$. Traditionally in the portfolio optimization literature, a portfolio has been represented by its weight vector $(w_1, w_2, \ldots, w_n)$ but it can also be equivalently represented in terms of its amount allocation vector $(x_1, x_2, \ldots, x_n)$.

In practice, the historical data is used to estimate the parameters in the optimization problem (6.1). Let r_{it} be the realization of the random variable r_i during the period t $(t = 1, 2, \ldots, T)$. We assume that r_{it} is available through the historical data and the expected value of r_i can be approximated by the average derived from the data. This gives

$$\mu_i = E(r_i) = \frac{1}{T} \sum_{t=1}^{T} r_{it}. \tag{6.2}$$

Also $w_L(x_1, x_2, \ldots, x_n)$ can be approximated by

$$E\left(\left|\sum_{i=1}^{n}(r_i-\mu_i)x_i\right|\right)=\frac{1}{T}\sum_{t=1}^{T}\left|\sum_{i=1}^{n}(r_{it}-\mu_i)x_i\right|. \tag{6.3}$$

In (6.3) it may be noted that, due to the absolute value function, the expression on the right hand side becomes a nonlinear and non smooth function of $(x_1, x_2, \ldots, x_n)$.

Using (6.2) and (6.3), problem (6.2) can be reformulated as

$$\begin{aligned}
&\text{Min} \quad \frac{1}{T}\sum_{t=1}^{T}\left|\sum_{i=1}^{n}(r_{it}-\mu_i)x_i\right| \\
&\text{subject to} \\
&\qquad \sum_{i=1}^{n}\mu_i x_i \geq \alpha M_0 \\
&\qquad \sum_{i=1}^{n} x_i = M_0 \\
&\qquad 0 \leq x_i \leq u_i \qquad (i=1,2,\ldots,n).
\end{aligned} \tag{6.4}$$

If we now denote $\left|\sum_{i=1}^{n}(r_{it}-\mu_i)x_i\right|$ by y_t and employ the definition of the absolute function then we get

$$\begin{aligned}
y_t &= \left|\sum_{i=1}^{n}(r_{it}-\mu_i)x_i\right| \\
&= Max\left(\sum_{i=1}^{n}(r_{it}-\mu_i)x_i, \quad -\sum_{i=1}^{n}(r_{it}-\mu_i)x_i\right).
\end{aligned} \tag{6.5}$$

Therefore problem (6.4) gets transformed to

$$
\begin{aligned}
&\text{Min} \quad \frac{1}{T}\sum_{t=1}^{T} y_t \\
&\text{subject to} \\
&\qquad \sum_{i=1}^{n}(r_{it}-\mu_i)x_i \le y_t \\
&\qquad -\sum_{i=1}^{n}(r_{it}-\mu_i)x_i \le y_t \\
&\qquad \sum_{i=1}^{n}\mu_i x_i \ge \alpha M_0 \\
&\qquad \sum_{i=1}^{n} x_i = M_0 \\
&\qquad 0 \le x_i \le u_i \qquad (i = 1,2,\ldots,n),
\end{aligned}
\tag{6.6}
$$

where the first two constraints in (6.6) follow from (6.5) and the definition of maximum.

Denoting $(r_{it} - \mu_i)$ by c_{it} $(i = 1,2,\ldots,n;\ t = 1,2,\ldots,T)$, we can rewrite (6.6) as

$$
\begin{aligned}
&\text{Min} \quad \frac{1}{T}\sum_{t=1}^{T} y_t \\
&\text{subject to} \\
&\qquad y_t - \sum_{i=1}^{n} c_{it}x_i \ge 0 \qquad (t = 1,2,\ldots,T) \\
&\qquad y_t + \sum_{i=1}^{n} c_{it}x_i \ge 0 \qquad (t = 1,2,\ldots,T) \\
&\qquad \sum_{i=1}^{n}\mu_i x_i \ge \alpha M_0 \\
&\qquad \sum_{i=1}^{n} x_i = M_0 \\
&\qquad 0 \le x_i \le u_i \qquad (i = 1,2,\ldots,n)\,.
\end{aligned}
\tag{6.7}
$$

Remark 6.3.1 *The formulation (6.4) is essentially nonlinear and non smooth. But since this nonlinearity and non smoothness occurs due to the presence of absolute value function only, it can be handled in a reasonably simple manner* *as explained above. The resulting linear programming problem (6.7) can be solved efficiently even when n is large.*

One obvious question at this stage is to enquire if there is any relationship between L_1 and L_2-risk models. The below given theorem answers this question.

Theorem 6.3.1 *Let $(r_1, r_2, \dots, r_n)$ be multivariate normally distributed. Then for a given portfolio $x = (x_1, x_2, \dots, x_n)$*

$$w_{L_1}(x) = \sqrt{\frac{2}{\pi}}\sigma(x),$$

where the standard deviation $\sigma(x)$ is given by

$$\sigma(x_1, x_2, \dots, x_n) = \sqrt{E\left[\left\{\sum_{i=1}^{n} r_i x_i - E\left(\sum_{i=1}^{n} r_i x_i\right)\right\}^2\right]}.$$

Proof. Let $(\sigma_{ij}) \in \mathbb{R}^{n\times n}$ be the variance-covariance matrix of $(r_1, r_2, \dots, r_n)$. Then under the given hypothesis, $\sum_{i=1}^{n} r_i x_i$ is normally distributed with mean $\sum_{i=1}^{n} r_i \mu_i$ and standard deviation

$$\sigma(x) = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j}.$$

Therefore

$$\begin{aligned} w_{L_1}(x) &= \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{\infty} |\, u \,| \exp\left(-\frac{u^2}{2\sigma^2(x)}\right) du \\ &= \frac{2}{\sqrt{2\pi}\sigma(x)} \int_{0}^{\infty} u \exp\left(-\frac{u^2}{2\sigma^2(x)}\right) du, \end{aligned}$$

which on substitution $(u^2/2\sigma^2) = s$ gives

$$w_{L_1}(x) = \sqrt{\frac{2}{\pi}}\,\frac{\sigma^2(x)}{\sigma(x)} \int_{0}^{\infty} e^{-s} ds = \sqrt{\frac{2}{\pi}}\sigma(x).$$

□

In view of Theorem 6.3.1, for the multivariate normal case, both L_1 and L_2-risk models are equivalent. In other words, Markowitz's mean-variance portfolio selection strategy will be the same as the one given by mean-absolute deviation selection strategy. Even in the case when normality assumption does not hold, through certain case studies, it has been shown that minimizing the L_1-risk produces portfolios which are comparable to Markowitz's mean-variance model which minimizes L_2-risk, i.e. standard deviation. However as one expects the variance of mean-absolute deviation portfolio is always at least as large as the corresponding mean-variance portfolio. But in actual applications, this difference is small.

In fact Konno and Yamazaki [80] applied both L_1 and L_2-risk models in Tokyo Stock Market by using historical data of 224 stocks in NIKKEL 225 index. They generated efficient frontiers and observed that the difference of the standard deviation of the optimal portfolio generated by L_2 and L_1-risk models is at most 10% for what ever value of α. Of course two frontiers will coincide if r_i's are multivariate normally distributed. Thus this difference can be largely attributed to the non normality of the data. Therefore irrespective of the distribution scenario, L_1-risk model provides a good alternative to Markowitz's L_2-risk model.

Advantages of the Formulation (6.7)

We now list some of the advantages of L_1-risk model (6.7) over the classical Markowitz's model which minimizes the L_2-risk.

(i) The formulation (6.7) is a linear programming problem and hence can be solved much more efficiently in comparison to quadratic programming problem which is obtained in Markowitz's model. This is particulary significant for portfolios having large number of assets.

(ii) We do not need to calculate the variance-covariance matrix to set up the L_1-risk model.

(iii) In the formulation (6.7) there are always (2T+2) constraints regardless of the number of assets included in the model. This allows to handle very large portfolios on a real time basis.

(iv) An optimal solution $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ of problem (6.7) contains at most (2T+2) positive components if $u_i = +\infty$ for $i = 1, 2, \ldots, n$. This means that an optimal portfolio consists of at most (2T+2) assets regardless of the size of n. Therefore we can use T as a control variable when we wish to restrict the number of assets in the portfolio.

as explained above. The resulting linear programming problem (6.7) ... efficiently even when n is large.

One obvious question at this stage is to enquire if there is any relation between L_1 and L_2-risk models. The below given theorem answers this question.

Theorem 6.3.1 *Let $(r_1, r_2, \ldots, r_n)$ be multivariate normally distributed. Then for a given portfolio $x = (x_1, x_2, \ldots, x_n)$*

$$w_{L_1}(x) = \sqrt{\frac{2}{\pi}}\sigma(x),$$

where the standard deviation $\sigma(x)$ is given by

$$\sigma(x_1, x_2, \ldots, x_n) = \sqrt{E\left[\left\{\sum_{i=1}^{n} r_i x_i - E\left(\sum_{i=1}^{n} r_i x_i\right)\right\}^2\right]}.$$

Proof. Let $(\sigma_{ij}) \in \mathbf{R}^{n\times n}$ be the variance-covariance matrix of $(r_1, r_2, \ldots, r_n)$. Then under the given hypothesis, $\sum_{i=1}^{n} r_i x_i$ is normally distributed with mean $\sum_{i=1}^{n} r_i \mu_i$ and standard deviation

$$\sigma(x) = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_{ij} x_i x_j}.$$

Therefore

$$\begin{aligned} w_{L_1}(x) &= \frac{1}{\sqrt{2\pi}\sigma(x)} \int_{-\infty}^{\infty} |\, u \,| \exp\left(-\frac{u^2}{2\sigma^2(x)}\right) du \\ &= \frac{2}{\sqrt{2\pi}\sigma(x)} \int_{0}^{\infty} u \; \exp\left(-\frac{u^2}{2\sigma^2(x)}\right) du\,, \end{aligned}$$

which on substitution $(u^2/2\sigma^2) = s$ gives

$$w_{L_1}(x) = \sqrt{\frac{2}{\pi}}\,\frac{\sigma^2(x)}{\sigma(x)} \int_{0}^{\infty} e^{-s} ds = \sqrt{\frac{2}{\pi}}\sigma(x).$$

□

In view of Theorem 6.3.1, for the multivariate normal case, both L_1 and L_2-risk models are equivalent. In other words, Markowitz's mean-variance portfolio