Building an optimal portfolio using a Mean-VaR framework

FLORENTIN ȘERBAN, MARIA VIORICA ȘTEFĂNESCU, SILVIA DEDU
Department of Mathematics
Academy of Economic Studies
Calea Dorobanțiilor 15-17, 0105724, Bucharest
ROMANIA
florentin.serban@csie.ase.ro, viorica.stefanescu@yahoo.com, silvia.dedu@csie.ase.ro

Abstract: - The area of finance has been a continuous source of challenging problems that have influenced research efforts on analytical and numerical solution methods for complex decision problems. In this paper we propose an original algorithm for portfolio optimization. We attack the problem in three stages: selecting assets, risk estimation, portfolio optimization. We select assets in the portfolio using principal components analysis in order to construct the initial portfolio, then we select from each of the classes obtained those assets that correspond to the minimum measure Value-at-Risk at a fixed probability level. Finally we solve the optimization problem. One numerical example is also presented for the sake of illustration.

Keywords: - estimation; optimization; principal components analysis; risk; selection of assets; Value at Risk

1 Introduction

In portfolio optimization a portfolio manager is faced with the problem to select from a usually very large set of assets offered in the stock market a subset of assets for investment following a specific objective with respect to (future) performance and risk. Once a portfolio is generated it has to be controlled and re-optimized over time due to the dynamic of asset markets and here another aspect.

Portfolio optimization is a classical research problem in Operations Research. Research started with the mean–variance portfolio selection model by Markowitz (1952) which from a mathematical programming point of view is a quadratic optimization problem with linear constraints. The mean refers to the average result among a set of possible scenarios, while the risk dimension describes the possible variation of the results under varying scenarios. In the Markowitz model the risk is measured by the variance of the outcome.

In the mean–risk setting the decision maker is faced with a two-objective situation: he/she wants to maximize the mean return and to minimize the risk at the same time. As for all multi-objective situations, there is in general no uniquely defined best decision, which is optimal in both dimensions and one has to seek for compromises. It is a typical multivariate problem: the only way to improve future returns is to increase the risk level that the decision maker is disposed to accept.

There are several assumptions and consequences behind the Markovitz mean/variance model, such as returns are normally distributed, so that mean and variance are sufficient to fully describe the portfolio return distribution function. But in some occurrences this assumption is not respected by data. For example, a big amount of research pointed out that real financial data are characterized by fat tails. In other words, the probabilities of incurring on extreme losses or gains are much higher than predicted by the normality assumption. These observations lead to new research directions on portfolio models. Therefore, the recent trend for portfolio selection is to use probabilistic [8], logarithmic [7], quantile [5, 9], and integral quantile [10] criteria. Value-at-risk (VaR) has become a standard measure used in financial risk management due to its conceptual simplicity, computational facility, and ready applicability [11]. The risk measures and properties that risk measures should satisfy have recently received considerable attention in the financial mathematics literature. Mathematically, a risk measure is a mapping from a class of random variables to the real line. Economically, a risk measure should capture the preferences of the decision-maker.

This study is dedicated to solve the problem of optimal portfolio consisting of risky stock and it aims to maximize expected return in terms of risk taking in the market risk measured by BET index. Our work is related to Basak and Shapiro [2], Ahn et. al [1] in that these studies also investigate an optimization framework that takes VaR into consideration.
We propose solving the problem in three stages: selecting assets, risk estimation and portfolio optimization. The selection of assets is realized by applying principal components analysis in order to discover similarities between the assets under consideration. We use PCA to reduce the number of features of assets to be taken into account for each asset. Brida and Risso [3] have applied clustering techniques to classify the stocks of Milan and Frankfurt Stock Exchanges using the Pearson correlation coefficient. Since the stock market in Romania is not mature enough we can not afford to consider only the closing price to determine the "distance" between stocks. In the second stage, we will present an approach to estimating risk using historical simulation method. Third stage focuses on the determination of both low-risk assets in each class and the optimal portfolio construction. At the end we solve a case study for stocks listed on the Bucharest Stock Exchange.

2 Financial ratios used in the stocks evaluation

Evaluation of the stocks will be performed, as usual, using two specific methods of financial analysis: fundamental analysis and technical analysis. Fundamental analysis attempts to determine a value closer to the reality of actions based on information on the company's financial situation, the area in which they operate, investments, property etc. The purpose of this analysis is to select those stocks which, at the time, the market price is lower than the value of the outcome of the analysis, thus creating the premises for future market to recognize the value and price to rise. This means that fundamental analysis attempts to predict the direction of stock price development of medium and long term from past and present achievements of the company and to estimate its future.

Technical analysis studies the evolution of the trading price, it assumes that all relevant information to the market is already included in price, except for natural disasters such shock events etc.. Investor psychology is captured very well to technical analysis.

Depending on access to information, the time for analysis and investment strategy chosen, each investor chooses the type of analysis that fits better. Thus speculators go long on technical analysis, long-term investors go on fundamental analysis. Ideally the two should be used together to confirm the purchase or sale signals that they offer.

We will present the financial indicators that we will use in our study:
- The BV (Accounting value) of a stock is calculated by dividing the total amount of equity in the company by the total number of stock issued by it and that are in circulation. The value of proper equity is determined by subtracting total liabilities from the company's total assets owned and it represents the "wealth of shareholders", that is what they remain with in the assumption that these assets would be valued and all debts would be paid
- The PBR (The price-to-book ratio) of a stock is calculated by dividing the company's current stock price by the book value per stock
- PSR indicator is calculated by dividing the current market value per stock of turnover in the last 4 consecutive quarters
- Evolution of price: to observe the price level at a given time we take into account the maximum price and minimum price achieved in the last 9 months

We used information on a total of 60 stocks representing stocks of Class I and II, traded on the Bucharest Stock Exchange on 01.08.2011. Then we considered only the stocks very volatile and for which it is possible to calculate the indicators mentioned and resulting in 40 stocks. We take into account several characteristics for each stock, we use data analysis techniques in order to process this vast amount of information.

We consider for each asset the values of five characteristics described above on 01.08.2011.
### 3 Stage of selection of assets

In the context of nowadays financial markets it is a huge amount of available financial data. It is therefore very difficult to make use of such an amount of information and to find basic patterns, relationships or trends in data. We apply data analysis techniques in order to discover information relevant to financial data, which will be useful during the selection of assets and decision making. Consider that we have collected information on a number $S$ of assets, each with $P$ features, which represent various financial ratios, still called variables. Denote by $y_i^j$ the $j$-th variable for action $i$. Multivariate data set will be represented by a matrix $Y = (y_i^j)_{i=1}^{S}{}_{j=1}^{P}$ and can be viewed as a set of $S$ points in a $P$-dimensional space.

Principal components analysis (PCA) is a useful technique for analyzing data to find patterns of data in a large-scale data space. PCA involves a mathematical procedure that transforms $P$ variables, usually correlated in a number of $p \leq P$ uncorrelated variables called principal components. After applying the PCA, each asset $i$ will be characterized by $p$ variables, represented by a set of parameters $y_i^1, y_i^2, \ldots, y_i^p$. Therefore, it is possible to form the arrays $Y_i = (y_i^1, y_i^2, \ldots, y_i^p)$, $i = 1, S$, which correspond to a set of $S$ assets. Suppose now that we obtained a data set $Y_i = (y_i^1, y_i^2, \ldots, y_i^p)$, $i = 1, S$. We then use clustering techniques in order to find similarities and differences between the actions under consideration. The idea of clustering is an assignment of the vectors $Y_1, Y_2, \ldots, Y_S$ in $T$ classes $C_1, C_2, \ldots, C_T$.

Once completed the selection of activities, we construct the initial portfolio by selecting low-risk asset in each class. We apply data analysis techniques to discover the similarities and differences between the stocks of the Bucharest Stock Exchange, using the package XLStat.

Figure 1 contains the tree resulted from PCA (dendogram). Dendogram usually begins with all assets as separate groups and shows a combination of mergers to a single root. Stocks belonging to the same cluster are similar in terms of features taken into account. In order to build a diversified portfolio, we first choose the number of clusters, which will be taken into account. We will then choose a stock from each group and we get the initial portfolio.
4 Phase estimation risk.

We evaluate the performance of an asset using expected future income, an indicator widely used in financial analysis. Denote by \( S_j(t) \) the closing price for an asset \( j \) at time \( t \). Expected future income attached to the time horizon \([t, t+1]\) is given by:

\[
R_j(t) = \ln S_j(t+1) - \ln S_j(t), \quad j \in \mathbb{I}, S.
\]

Similarly, we define the loss random variable, the variable \( L_j \), for asset \( j \) for \([t, t+1]\) as:

\[
L_j(t) = -R_j(t) = \ln S_j(t) - \ln S_j(t+1), \quad j \in \mathbb{I}, S.
\]

Using Rockafellar et al.[14] define the risk measure VaR corresponding loss random variable \( L_j \).

Probability of \( L_j \) not to exceed a threshold \( z \in \mathbb{R} \) is \( G_{L_j}(z) = P(L_j \leq z) \).

Value at risk of loss random variable \( L_j \) associated with the value of asset \( j \) income and corresponding probability level \( \alpha \in (0,1) \) is:

\[
VaR_{\alpha}(L_j) = \min \left\{ z \in \mathbb{R} \left| G_{L_j}(z) \geq 1 - \alpha \right\} \right. \text{ or } P(X > VaR) = \alpha. \]

If \( G_{L_j} \) is strictly increasing and continuous, \( VaR_{\alpha}(L_j) \) is the unique solution of equation \( G_{L_j}(z) = 1 - \alpha \) then \( VaR_{\alpha}(L_j) = G_{L_j}^{-1}(1 - \alpha) \).

One of the most frequently used methods for estimating the risk is the historical simulation method. This risk assessment method is useful if empirical evidence indicates that the random variables in question may not be well approximated by normal distribution or if we are not able to make assumptions on the distribution. Historical simulation method calculates the value of a hypothetical changes in the current portfolio, according to historical changes in risk factors. The great advantage of this method is that it makes no assumption of probability distribution, using the empirical distribution obtained from analysis of past data. Disadvantage of this method is that it predicts the future development based on historical data, which could lead to inaccurate estimates if the trend of the past no longer corresponds. If \( L_j \) is the loss random variable and \( \hat{G}_n \) is empirical distribution function of \( L_j \) and \( \alpha \in (0,1) \) a fixed level of probability, then \( \hat{G}_n(z) = \frac{1}{n} \sum_{i=1}^{n} I_{(L_i \leq z)} \).

We can prove that

\[
VaR(L_j) = \min \left\{ z \in \mathbb{R} \left| \frac{1}{n} \sum_{i=1}^{n} I_{(L_i \leq z)} \geq 1 - \alpha \right\} . \right.
\]

We used the closing price values daily for each share, corresponding to a time horizon of 50 days to measure VaR for each stock. We used the data available on the Bucharest Stock Exchange from 15 may-30 June 2011. The following tables contain values of VaR for each stock and three levels of probability values.
5 Optimization portfolio phase

Once completed the phase of grouping the assets in T classes by the existing similarities, we focus on the selection of the assets of each class to have a minimal risk. Consider the loss random variable corresponding to each asset in each obtained class $T_i C_i$, $i = 1, T$. We obtain an initial portfolio comprising a wide range of actions with minimal risk.

As a consequence of applying the technique of selection, we selected a portfolio of 5 stocks, each of them representing the minimum risk stock class corresponding VaR measured probability level 0.99: Tlv, Bio, Sno, Sif5, Tgn. Then we will try to determine what percentage is the optimal composition of capital that needs to be invested in each of the assets under consideration, so that at the end of the investment we have a maximum return on investment.

Thus, $T$ is a set of stocks, with stock $j$ that leads to expected income $R_j$, $j = 1, T$; expected income of portfolio is $R_x = \sum_{j=1}^{T} x_j R_j$. The model to be solved is:

$$\max \sum_{j=1}^{T} x_j R_j, \text{with VaR}_\alpha (L_x) \leq \nu_0, \nu_0 \text{ is parameter.}$$

In these conditions, the optimization problem is:

$$\max f = 0.0935 x_1 + 0.0238 x_2 + 0.0338 x_3 + 0.1437 x_4 + 0.1716 x_5$$

$$0.0259 x_1 + 0.0434 x_2 + 0.0368 x_3 + 0.0295 x_4 + 0.0327 x_5 \leq \nu_0$$

$$\sum_{i=1}^{5} x_i = 1$$

$$x_i \geq 0, i = 1, 5$$

- We would like that $VaR_\alpha (L_x) \leq \nu_0$, where $\nu_0$ is the capital market risk, calculated as measure of BET index risk measured by VaR, ie

The calculations using table 2 for the 0.99 level of probability lead to $\nu_0 \approx 0.035$
- Return of stock is: $P_t - P_{t-1}$ and the problem is:
\[
\max f = 0.0935 x_1 + 0.0238 x_2 + 0.0338 x_3 + 0.1437 x_4 + 0.1716 x_5 \\
0.0259 x_1 + 0.0434 x_2 + 0.0368 x_3 + 0.0295 x_4 + 0.0327 x_5 \leq 0.035 \\
\sum_{i=1}^{5} x_i = 1 \\
x_i \geq 0, \; i = 1,5
\]

Solving this problem leads to the optimal solution

Table 3: Optimal solution

<table>
<thead>
<tr>
<th>$(r_0)$</th>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_4$</th>
<th>$x_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.026</td>
<td>0.972</td>
<td>0</td>
<td>0</td>
<td>0.028</td>
<td>0</td>
</tr>
<tr>
<td>0.027</td>
<td>0.694</td>
<td>0</td>
<td>0</td>
<td>0.306</td>
<td>0</td>
</tr>
<tr>
<td>0.028</td>
<td>0.417</td>
<td>0</td>
<td>0</td>
<td>0.583</td>
<td>0</td>
</tr>
<tr>
<td>0.029</td>
<td>0.139</td>
<td>0</td>
<td>0</td>
<td>0.861</td>
<td>0</td>
</tr>
<tr>
<td>0.030</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.844</td>
<td>0.156</td>
</tr>
<tr>
<td>0.031</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.531</td>
<td>0.469</td>
</tr>
<tr>
<td>0.032</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.219</td>
<td>0.781</td>
</tr>
<tr>
<td>0.033</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.996</td>
</tr>
<tr>
<td>0.034</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.982</td>
</tr>
<tr>
<td>0.035</td>
<td>0.031</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.969</td>
</tr>
</tbody>
</table>

6 Conclusions

The proposed methodology has been highlighted by a case of study for the stocks listed on BSE. To have the risk assumed by the investor under the risk of the capital market (measured by the BET index), applying the algorithm presented in this study, we get the combinations presented in Table 3. Therefore:
- at the level of the assumed risk of 0.035 (the level of the risk of the BET index) the optimal portfolio obtained is formed by 97% Tgn and 3% Bio;
- the proposed problem has optimal solutions only if the level of the risk lies in the interval (0.026, 0.035);
- we cannot construct any optimal portfolio if the assumed risk is under 0.026;
- in order to have a minimal assumed risk, the optimal portfolio is: 97.2% stocks Bio and 2.8% stocks Tlv.
- we can see that by a performant comparison of the portfolios, the best is the STAT portfolio, which shows the efficacy of our methods.

Acknowledgment

This work was supported from the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013, project number POSDRU/89/1.5/S/59184 „Performance and excellence in postdoctoral research in Romanian economics science domain”

References

[16] [http://bvb.ro/](http://bvb.ro/)