Lecturer Florentin SERBAN, PhD Professor Viorica STEFANESCU, PhD Department of Mathematics The Bucharest Academy of Economic Studies Professor Massimiliano FERRARA, PhD Department of Mathematics University of Reggio Calabria, Italy

# PORTFOLIO OPTIMIZATION AND BUILDING OF ITS EFFICIENT FRONTIER

**Abstract.** This paper presents the description of the efficient frontier for a portfolio made of three assets. We use data analysis to obtain three clusters, then, we estimate the risk of each asset corresponding to each class we obtained. Thus, we get the best three assets among the ones we analyzed and for which we will construct the efficient frontier. The originality of our paper consists in the combination of classification theory and risk estimation theory to determine the best assets. To illustrate the efficiency of the method we used, we present a case study which makes reference to the stocks listed at BSE. We construct the efficient frontier based on the existent correlation of the best analyzed stocks that we obtained by data analyses (for classification), and by the evaluation of the loss repartition (for risk estimation)

*Keywords: risk; selection of assets; principal components analysis; optimization; efficient frontier.* 

## JEL CLASSIFICATION : C02, C61, G11.

## 1. Building an optimal portfolio consisting of n assets

#### **1.1. Introduction**

Financial assets portfolio optimization is an important area, which developed the theory of Markowitz's mean-variation and the expected utility theory. Meandispersion theory has some limitations and can be applied successfully only if the expected returns of financial securities are normally distributed random variables. However, the literature contradicts the hypothesis of normality for the expected returns and it is a strong argument against the use of the mean-variation techniques, which is why they introduced new measures of risk. Value-at-Risk (VaR) is a measure of risk, which plays an important role in investment, risk management and regulatory control of financial institutions. Basel II has incorporated the concept of VaR, and encourages banks to use VaR for the daily risk management. Since in most cases the distribution of

random variable risk is not known, a method of evaluation or estimation of VaR is required. The idea to obtain clusters that characterize a set of assets can be found also in Kaski et al, 2009, Stefanescu et al, 2008: pp.109-122, and Mantegna, 1999. Methodology based on clustering techniques is a useful tool for understanding and detecting the structure and the hierarchy in the financial data. These methods were successfully applied to analyze the stock markets and exchange. Brida and Risso (2007a, 2007b) applied clustering techniques to classify the stocks of Milan and Frankfurt stock exchanges using the Pearson correlation coefficient. This study is dedicated to solve the problem of optimal portfolio consisting of risky shares and it aims to maximize expected return in terms of risk taking in the market risk measured by BET.

We propose solving the problem in two stages: selecting assets, risk estimation a. 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. In the second stage, we will present an approach to estimating risk using historical simulation method.. At the end we solve a case study for stocks listed on the Bucharest Stock Exchange.

## **1.2. 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 N of assets, each with P features, which represent various financial ratios, still called variables. Denote by  $x_i^j$  the j-th variable for action i. Multivariate data set will be represented by a matrix  $X = (x_i^j)_{\substack{i=\overline{1,N}\\i=1}}$  and can be viewed as a set of N 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  $p \le P$  uncorrelated variables called principal components. After applying the PCA, each asset *i* will be characterized by *p* variables, represented by a set of parameters  $x_i^1, x_i^2, ..., x_i^p$  therefore, it is possible to form the arrays  $X_i = (x_i^1, x_i^2, ..., x_i^p)$ ,  $i = \overline{1, N}$ , which correspond to a set of S assets. Suppose now that we obtained a data set  $X_i = (x_i^1, x_i^2, ..., x_i^p)$ ,  $i = \overline{1, N}$ . 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  $X_1, X_2, ..., X_N$  in T classes  $C_1, C_2, ..., C_T$ . Once completed the selection of activities, we construct the initial portfolio by selecting low-risk asset in each class.

#### **1.3.** 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:  $R_j(t) = \ln P_j(t+1) - \ln P_j(t), \ j \in \overline{1, N}$ .

Similarly, we define the loss random variable, the variable  $L_j$ , for asset j for [t, t+1]as :  $L_j(t) = -R_j(t) = \ln P_j(t) - \ln P_j(t+1)$ ,  $j \in \overline{1, N}$ . Using Rockafellar et al., 2000, define the risk measure VaR corresponding loss random variable  $L_j$ . Probability of  $L_j$  not to exceed a threshold  $z \in \mathbf{R}$  is  $G_{L_j}(z) = P(L_j \le z)$ .

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:  $VaR_{\alpha}(L_j) = \min\{z \in \mathbf{R} \mid G_{L_j}(z) \ge 1-\alpha\}$  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  $L_i$  and  $\alpha \in (0,1)$  a fixed

level of probability, then 
$$\hat{G}_n(z) = \frac{1}{n} \sum_{i=1}^n F_{\{L_j \le z\}}$$
. We can prove that  $VaR(L_j) = \min\left\{z \in \mathbf{R} \mid \frac{1}{n} \sum_{i=1}^n F_{\{L_j \le z\}} \ge 1 - \alpha\right\}.$ 

## 2. The efficient frontier of a portfolio composed of three assets

Assume that on the market there are 3 risky actives

The active ",*i*" has percentage of profitableness  $r_i$ , where:

The average  $M(r_i) = \mu_i$ 

- Standard deviation of  $r_i = \sigma_i$
- For  $i, j = \overline{1.3}$ ;  $i \neq j$ , one knows the covariance  $\sigma_{ii}$
- Total correlation matrix is  $V = \|\sigma_{ij}\|_{i=1,3}$

We denote by  $x_i(i = \overline{1.3})$  -the weight of the active  $A_i$  within the portfolio.

The random variable which gives the profitableness of the portfolio:  $r_p = \sum x_i \cdot r_i$ Definition 1: A portfolio is called "profitable" if:

Among all the portfolios with the same standard deviation of the profitableness, has the best average.

*Definition 2* : A portfolio is called "efficient" if , starting from a set of shares and taking all the linear combination of the titles within the portfolio, we are looking for those titles which dominates other (profitable) titles

*Definition 3*: The set of profitable portfolios is called "the frontier of profitableness" (Markowitz:, 1990).

Consider that the sell of the live actives is not permitted (this means that the weights  $x_{1,2,3}$  of the actives within the portfolio are strictly positive) and that the risks

 $r_{1,2,3}$  of the actives have the averages  $\mu_{1,2,3}$  and deviations  $\sigma_{1,2,3}$ .

Therefore, the weights  $\{x_i\}_{i=\overline{1,3}}$  satisfy the relations  $x_i > 0$ ;  $i = \overline{1.3}$  and  $\sum_i x_i = 1$ 

$$m_p = \sum_i x_i \cdot \mu_i$$
$$D^2(r_p) = \sigma_p^2 \Leftrightarrow \sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n x_i \cdot x_j \cdot \sigma_{ij}$$

As a measurement of the risk we use the standard deviation of the profitableness of the portfolio.

In order to determine the optimal  $\{x_i\}_{i=\overline{1,n}}$ , we use the nonlinear model:

$$\begin{pmatrix} (\min)\sigma_p^2 = \sum_{i=1}^3 \sum_{j=1}^3 x_i \cdot x_j \cdot \sigma_{ij} \\ \sum_i x_i = 1 \\ \sum_i x_i \cdot \mu_i = m_p \end{cases}, x_i > 0; i = \overline{1.3}$$

The profitableness frontier of the portfolio in the  $(\sigma_p^2; m_p)$  is a parabola.



## **Comments**

1) The minimal risk achieved in the point (C) is not zero.

2) Only the segment (CA) is a profitableness frontier : a prudent investor cannot accept the segment (BC) because any portfolio on the segment (BC) is strictly dominated by a portfolio on the segment (CA) which has the same risk but a better profitableness.

## **3.** Application of optimization of a portfolio of stocks listed on BV B **3.1.** Financial ratios used in the actions evaluation

Evaluation of the actions 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 actions 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 raise.

This means that fundamental analysis attempts to predict the direction of share price development of medium and long term from past and present achievements of the company and to estimate its future. So, fundamental analysis relies on a direct cause-effect relationship between the economic value of a stock and its market price developments. *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

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 analyses, 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 some of the most important financial indicators that we will use in our study.

*PER indicator (net income per share)* is calculated by dividing the current market price to the value of net profit per share for the past four consecutive quarters,
 *The P / BV (book value of shares)* is calculated by dividing the current trading price to book value per share determined according to the latest financial reporting, accounting value of a share is calculated by dividing the total equity value of the company to the total of it shares issued and outstanding; equity value is determined by deducting total liabilities from total assets owned company and is "shareholder wealth", which is what remains to be recovered if the assets and liabilities would be paid.

- *The ratio of value traded in last 52 weeks and market capitalization*, the report shows the liquidity action

- *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 6 months

*-Divy index* measures the performance of dividend and is calculated as the ratio between the amount of the dividend and book value or market value of the action. Divy index assesses the efficiency of investment in an asset.

We used information on a total of 71 shares representing shares of Class I and II, traded on the Bucharest Stock Exchange. We selected only the actions characterized by a high value of the coefficient of volatility and we obtained a smaller set of 48 stocks. Then we considered only the actions for which it is possible to calculate all the indicators mentioned resulting in 39 stocks. The aim of our study is to find similarities and differences between the current analyses and build a diversified portfolio. Since the Bucharest Stock Exchange is not mature enough, we can not afford to use a single financial index, such as, for example, the closing price. So we take into account several characteristics for each asset, we use data analysis techniques in order to process this vast amount of information. We consider the values of the seven financial indices for each asset. Table 1 lists, for each of the 39 analyzed stocks , the values of the six features; we used the data available on the Bucharest Stock Exchange on 30 November 2010.

Portfolio Optimization and Building of its Efficient Frontier

## The value of the 7 features

## Table 1

No	Symbol	P/BV	PER	DIV Y	Min/ P	Max/P	TrV/Cap
1	ALR	1.39	16.61	6.70	0.89	1.41	0.005
2	ALT	0.31	13.26	0.00	0.86	1.43	0.38
3	ALU	0.72	26.88	5.44	0.73	1.9	0.12
4	ARS	1.52	11.79	6.94	0.71	1.16	0.05
5	ATB	1.04	14.63	3.00	0.95	1.27	0.1
6	AZO	0.62	2.68	0.00	0.87	1.17	0.21
7	BCM	0.37	18.51	7.83	0.71	1	0.09
8	BIO	1.55	10.63	0.00	0.83	1	0.4
9	BRD	1.77	14.1	2.33	0.87	1.15	0.07
10	BRK	0.84	13.8	0.00	0.86	1.26	1.09
11	BRM	0.62	6.17	9.09	0.78	1.5	0.11
12	BVB	3.06	27.4	2.84	0.79	1.11	0.39
13	CEON	0.41	26.66	0.00	0.80	1.25	0.4
14	CGC	0.22	117.8	0.00	1	1.8	0.06
15	CMP	0.34	14.1	0.00	0.88	1.04	0.13
16	COMI	0.97	12.82	0.00	1	1.95	0.97
17	DAFR	0.71	16.86	0.00	0.89	1.16	0.76
18	EFO	0.08	6.18	0.00	0.93	1.2	0.08
19	ELJ	0.32	2.28	0.07	0.68	1.17	0.009
20	OIL	0.86	41.39	0.26	0.72	1.1	0.1
21	PREH	0.44	46.65	4.86	0.97	1.66	0.04
22	PTR	0.77	6.66	4.92	1	1.91	0.41
	RMA						
23	Н	0.84	7.04	1.48	0.85	1.37	0.05
24	SCD	1.40	7.67	0.00	0.73	1.01	0.07
25	SIF1	1.08	12.31	5.1	0.98	1.17	0.55
26	SIF2	1.16	5.97	5.88	0.95	1.24	1.21
27	SIF3	0.85	22.71	5.88	0.91	1.45	0.6
28	SIF4	0.35	7.19	6.78	0.92	1.3	0.4
29	SIF5	1.12	11.62	12.8	0.9	1.40	1.68
30	SNO	0.49	26.12	10	1	1.55	0.08
31	SNP	1.32	8.95	0.00	0.77	1.02	0.01
32	SOCP	1.08	57.68	0.00	0.57	1.26	0.38
33	SPCU	0.60	128	3.19	1	1.16	0.2
34	SRT	0.15	16.74	0.00	1	2.19	0.35

35	TGN	1.30	7.82	4.98	0.79	1.02	0.05
36	TLV	0.87	20.2	2.97	0.99	1.67	0.3
37	TRP	0.86	20.08	0.00	1	1.32	0.17
38	TUFE	0.48	32.55	0.00	1	1.1	0.03
39	VNC	0.70	17.77	0.00	0.98	1.08	0.07

## 3.2 Principal components analysis

We apply data analysis techniques to discover the similarities and differences between the stocks of the Bucharest Stock Exchange, using the package StatistiXL 1.8. Figure 1 contains the tree resulted from PCA Dendrogram 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 ( for our study, we chose 3), which will be taken into account. We will then choose a stock from each group.



Dendrogram

**Comments:** we observe the 3 classes in which the shares were grouped. Those are the shares grouped around Bio, the ones around Tlv and the share Tgn which is associated much later to the already formed classes.

## **3.3.** Portfolio optimization

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 June 2010-30 July 2010. The following tables contain values of VaR for each stock and three levels of probability values.

Class 1	0.90	0.95	0.99
ALR	0.0139	0.0171	0.0519
ALT	0.0290	0.0371	0.0653
ALU	0.0236	0.0317	0.0529
ATB	0.0221	0.0293	0.0330
BRD	0.0195	0.0231	0.0361
CEON	0.0287	0.0455	0.0995
CMP	0.0228	0.0422	0.0701
COMI	0.0282	0.0313	0.0420
ELJ	0.0531	0.0742	0.0969
PREH	0.0568	0.0807	0.1476
RMAH	0.0465	0.0550	0.1003
SCD	0.0248	0.0471	0.0750
SNP	0.0201	0.0248	0.0325
SOCP	0.0317	0.0485	0.0584
SPCU	0.0418	0.0585	0.2624
SRT	0.0534	0.0687	0.0748
TEL	0.0247	0.0347	0.0492
TLV	0.0134	0.0204	0.0259
TRP	0.0415	0.0705	0.1025
VNC	0.0348	0.0383	0.0584
SIF1	0.0276	0.0317	0.0406
SIF2	0.0225	0.0270	0.0329
SIF3	0.0255	0.0305	0.0343
SIF4	0.0237	0.0280	0.0388
SIF5	0.0242	0.0260	0.0295
Class 3	0.90	0.95	0.99
TGN	0.0167	0.0221	0.0327

Class 2	0.90	0.95	0.99
AZO	0.0265	0.0405	0.0669
BIO	0.0235	0.0372	0.0434
BRK	0.0223	0.0424	0.0486
BVB	0.0270	0.0299	0.0676
CGC	0.0512	0.0747	0.1289
DAFR	0.0328	0.0400	0.0468
EFO	0.0531	0.0742	0.0969
ARS	0.0721	0.0891	0.1325
BCM	0.0460	0.0562	0.0815
BRM	0.0572	0.0650	0.1074
PTR	0.0348	0.0448	0.0566
TUFE	0.0333	0.0388	0.0558

## 3.4 Construction of an optimal portfolio made of 3 stocks

We start from the classes we formed above and we choose from each of them the stock which has minimal VaR for the probability 0.99; we get the three stocks which form the portfolio to which we can construct the frontier of profitableness: TLV, BIO and TGN

## 3.5 The efficient frontier of profitableness for the portfolio made of 3 stocks

We selected the closing prices for the period 9.11-10.12.2011 for the 3 selected stocks: TLV and BIO and TGN. We obtained the values:

TLV : 1.27; 1.29; 1.26; 1.25; 1.3; 1.3; 1.3; 1.23; 1.23; 1.23; 1.22; 1.23; 1.24; 1.24; 1.25; 1.27; 1.24; 1.24; 1.26; 1.26; 1.27

BIO: 0.185; 0.187; 0.187; 0.187; 0.188; 0.19; 0.189; 0.189; 0.188; 0.188; 0.188; 0.188; 0.188; 0.188; 0.189; 0.193; 0.2; 0.196; 0.196; 0.201; 0.203; 0.202; 0.203

TGN :261.2; 262.2; 257; 258; 263.7; 263.1; 262; 263; 263.5; 264; 263.1; 263; 262.1; 262; 262; 263; 262.8; 263.5; 263; 262; 262.5

We will transform the prices in annual profitableness by the formula (final price-initial price)x360/ initial price

We get the following repartition of the profitableness repartition :

$X_{TLV}$ :	$-28.8$ $\frac{1}{20}$	$-18$ $\frac{3}{20}$	$-10.8$ $\frac{4}{20}$	$\frac{-8.3}{20}$	$0$ $\frac{5}{20}$	$5.8$ $\frac{2}{20}$	$\frac{14.4}{20}$	$\frac{28.8}{\frac{1}{20}}\right).$	
$X_{BIO}$ :	$-7.2$ $\frac{1}{20}$	$-1.8$ $\frac{3}{20}$	$0$ $\frac{7}{20}$	$\frac{1.8}{\frac{4}{20}}$	$\frac{3.6}{\frac{2}{20}}$	$7.2$ $\frac{1}{20}$	$\frac{10.8}{20}$	$\left(\frac{14.4}{20}\right).$	
$X_{TGN}$ :	$-7.1$ $\frac{1}{20}$ ly the f	$-1.5$ $\frac{4}{20}$ formula	-0.8 $\frac{2}{20}$ ae speci	$-0.4$ $\frac{2}{20}$ fit to s	$\begin{array}{c} 4 & 0 \\  & \frac{1}{20} \\  tatistic \end{array}$	$ \begin{array}{r} 0.8 \\ \hline 6 \\ \hline 20 \\ \hline s \text{ and} \end{array} $	$\frac{1.5}{20}$ we get	$\frac{1}{20}$	

 $\mu_1 = 3.1$  and deviation  $\sigma_1 = 16$ 

 $\mu_2 = 1.7$  and deviation  $\sigma_2 = 4.7$ 

 $\mu_3 = 3.9$  and deviation  $\sigma_3 = 4.6$ 

 $cov_{12} = 12$ ;  $cov_{13} = 17$ ;  $cov_{23} = -5$ 

Let x,y,z be the weights of the 3 stocks. We get, according to the relations from section 2:  $m_p = 3.1 \text{ x} + 1.7 \text{ y} + 3.9 \text{ z}$ 

 $\sigma_p^2 = 256 x^2 + 22 y^2 + 21 z^2 + 2 x 12 xy + 2 x 17 xz - 2 x 5 yz$ 

We therefore have the following problem of nonlinear optimization: Min  $\sigma_p^2 = 256 x^2 + 22 y^2 + 21 z^2 + 24 xy + 34 xz - 10 yz$ 

 $\begin{cases} 3, x + 1.7 y + 3.9 z = m \\ x + y + z = 1 \end{cases}$ (1) (2)

$$x, y, z > 0$$
 (3)

We replace x and y in relation 1 and 2 in the goal function and we get  $\sigma_p^2 = 881z^2 - 2(116m - 672)z + 148m^2 + 274m + 1038$ 

Minimum of this function is achieved for  $z = \frac{(116m - 672)}{881} = 0.1m - 0.7$ 

We replace z in the goal function and we get  $\sigma_p^2 = 133 \text{ m}_p^2 - 168 \text{ m}_p + 528$ 

In the profitableness-risk plane the previous relation represents the frontier of profitableness for the portfolio made of the BIO, TLV and TGN. Graphically, we get the below representation with the remark that the efficient frontier is just the curve AC



#### 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

[1]Armeanu, D., Balu, F. (2008), Testing the Efficiency of Markowitz Model on Bucharest Stock Exchange. Economic Computation and Economic Cybernetics Studies and Research, 42 (1-2), ASE Publishing House, Bucharest;
[2]Brida, J. G., Risso, W. (2007), Dynamics and Structure of the Italian Companies, <u>http://ssrn.com/</u>;

[3]Brida, J. G., Risso, W. (2007), *Hierarchical Structure of the German Stock Market.* <u>http://ssrn.com/;</u>

[4]Dedu, S., Fulga, C. (2011), Value-at-risk. Estimation comparative aproach with application to optimization problems. Economic Computation and Economic Cybernetics Studies and Research, 45 (1), 127-142, ASE Publishing House, Bucharest
[5]Fulga, C., Dedu, S. (2009), Modeling Risk Using Risk Measures VaR and Cvar with Applications in Portfolio Management. Studii si cercetari de calcul economic si cibernetica economica, 43 (1-2), 99-112, ASE Publishing House, Bucharest;
[6]Fulga, C., Dedu, S., Şerban, F. (2009), Portfolio Optimization with Prior Stock Selection. Economic Computation and Economic Cybernetics Studies and Research, 43 (4), 157-172, ASE Publishing House, Bucharest;

[7] Larsen, N., Mauser, H., Uryasev, S. (2002), Algorithms for Optimization of Value-at-Risk. Financial Engineering, E-Commerce and Supply Chain, 129-157;
[8] Rockafellar, T., Uryasev, S.(2000), Optimization of Conditional Value-at-Risk. Journal of Risk, 2;

[9]**Stefānescu, M.V, Serban, F., Buşu, M., Ferrara, M. (2010)**, *Portfolio Optimization Using Classification and Functional Data Analysis Techniques*. *Economic Computation and Economic Cybernetics Studies and Research, 44 (3)*, *93-108*, ASE Publishing House, Bucharest.