Application of multivariate lambda distribution within the portfolio selection model

Mário Pčolár[1]

[1] University of Economics in Bratislava, Department of Operations Research and Econometrics, Dolnozemská cesta 1/b, 852 35 Bratislava, Slovakia

mario.pcolar@euba.sk

https://doi.org/10.53465/EDAMBA.2021.9788022549301.383-389

Abstract. The article deals with the application of the resampling procedure using multivariate lambda distribution within the process of optimization of portfolio selection models. The aim of the resampling procedure is to achieve portfolios that provide better quality results on out-of-sample data compared to the traditional optimization-based approach using estimates from historical data. In this paper, we deal with the application of the resampling procedure on daily data of 30 assets within the model of portfolio selection in the space of expected return and CVaR (Conditional Value at Risk). We are dealing with the application of two approaches, an approach based on the assumption of normal distribution of data using multivariate normal distribution for data generation and a procedure using data generation from multivariate generalized lambda distribution.

Keywords: CVaR optimization, multivariate distribution, resampling.

JEL classification: G11, G17

1 Introduction

In the past, the application of the resampling procedure was dealt with by several authors, most of whom dealt with the application within the Markowitz model using monthly or weekly data [2], [5]. Only a small number of contributions dealt with the application on other models of portfolio selection or on applications in the case of using daily data. The aim of this paper is to apply the resampling procedure within the model of portfolio selection in the space of expected return and CVaR using daily data. The paper also deals with the application of two modifications of this procedure, namely the procedure using multivariate normal distribution and the procedure using multivariate generalized lambda distribution (GLD). The Gld distribution was selected based on previous research [8]. In the paper we use the modification of the CVaR model to the task of linear programming, originally presented in [10]. Within the computational experiments we use FKML parameterization of the GLD distribution, a detailed overview of the issues concerning the GLD distribution as well as the formulation of the quantile function can be found, for example, in [1].

2 Resampling procedure

The procedure is generally based on the use of multivariate distribution on data generation and the Monte Carlo method, with a selected portfolio selection optimization model applied to each data simulation. The result of these procedure is a set of efficient frontiers quantified from individual random realizations of data, while the resulting frontier is obtained by averaging the weights of individual statistically equivalent portfolios.

Generating data from a multivariate normal distribution is a fairly well-known procedure that is currently programmed in most statistical software. In our paper, we use the rmvnorm() function contained in the Mvtnorm package in R [3]. In the case of generating data from a multivariate generalized lambda distribution, it is a bit more complicated. In the paper, we use the procedure of generating data from a multivariate non-normal distribution given in [4]. The application of this procedure requires that the distributions of the individual components of the random variable X must be known in the form of quantile functions and the correlation is available as a correlation matrix R_X using rank-based correlation (e.g. Spearman's correlation coefficient). Subsequently, the procedure consists of the following steps [4]:

1. Transform matrix R_x to matrix C_z applying statement:
$$C_Z = 2\sin\left[\frac{\pi}{6}\ R_X\right]$$
2. Generate data samples of m – dimensional normal distribution with correlation matrix C_z
3. Transform normal components into the components of uniform distribution by applying the distribution function (CDF) of the normalized normal distribution $U_i = \Phi(Z_i)$ such that $U_i \sim U(0,1)$
4. Quantify values of X_i using a given quantile function of individual components, $X_i = Q_{X_i}(U_i)$

The advantage of such procedure is considerable flexibility in the choice of the assumed distribution to the point that the quantile function must be known for the selected distribution.

Procedure of resampling within portfolio selection models was introduced by [7] and used in contributions like [9], [12]. We can classify it into the category of heuristics to solve the problem of portfolio selection [11]. Such a procedure reduces the problem of error maximization, for a more detailed description of this problem we recommend the paper [6]. The resulting portfolios of this procedure are more diversified compared to the traditional approach. The resampling procedure consists of the following steps:

- Estimate the parameters of the assumed probability distribution from historical data and estimate the covariance matrix.
- Than generate a vector of random realizations from a multivariate marginal probability distribution, using an estimated covariance matrix. The length of the generated interval is traditionally the same as the number of observations of historical data that we used in the previous step.
- The generated sample of data will then be used as input data for the portfolio selection model for estimating the efficient frontier. Save the values of weights for M evenly distributed portfolios at the efficient frontier by rank (from portfolio with minimal risk to portfolio with maximal return).
- Repeat the previous 2 steps many times, then average the weights of portfolios that share the same rank within the individual simulations.

3 Experiment results

Calculation of experiment are perform using daily data of closing positions of 30 assets, from 01.01.2012 to 31.12.2020. The assets consisted of DJIA components. We quantified daily returns as so-called logarithmic returns, i.e. the first difference of the natural logarithm of individual observations for individual assets. We use series of daily returns to estimating the parameters of the normal distribution and the parameters of the generalized lambda distribution, which we will use subsequently in the resampling procedure.

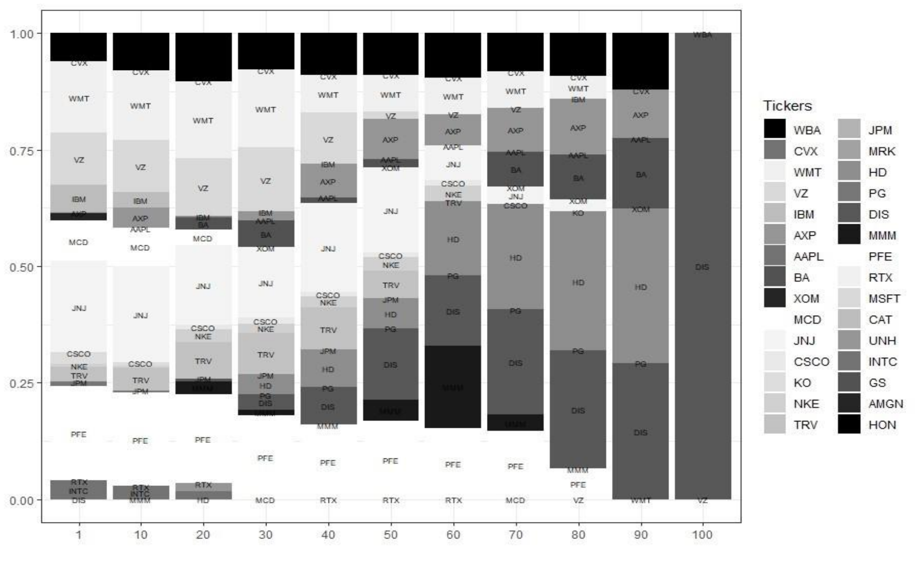

Fig. 16. Distribution of weights within selected portfolios at the efficient frontier quantified using the CVaR model and using of a "classical" approach.

To estimate the parameters of the distribution, we used the maximum likelihood estimation method, where when estimating the parameters, we consider a data sample for a period of two years, from 01.01.2012 to 31.12.2013. In Monte Carlo simulations, we performed 500 simulations, and in portfolio optimization using a CVaR risk model, we generated 100 evenly distributed portfolios at the efficient frontier.

In a computational experiment, we compare portfolios quantified using a "classical" approach using estimates from historical data compared to two resampling procedures, where in one case we are generating data from a multivariate normal distribution (Norm_mult) and in the other case we are generating data from a multivariate generalized lambda distribution (Gld_mult). We compare the performance of portfolios on a out-of-sample data, so without the data which we used to estimate the parameters of individual models. Specifically, it is a data sample from 01.01.2014 to 31.12.2020. As part of the experiment, we consider the scenario of an investor who invests in individual assets in accordance with selected constructed portfolios. There are three investment scenarios, a one-year investment horizon, a three-year investment horizon and a six-year investment horizon. An investor cannot sell the assets in which he has invested for the duration of the investment horizon. Investor can sell the assets only during the year following the investment horizon, but for a maximum period of one year. Specifically, in years 2015, 2017, 2020. In the experiment, we abstract from stock exchange fees and additional costs associated with the sale and purchase of assets.

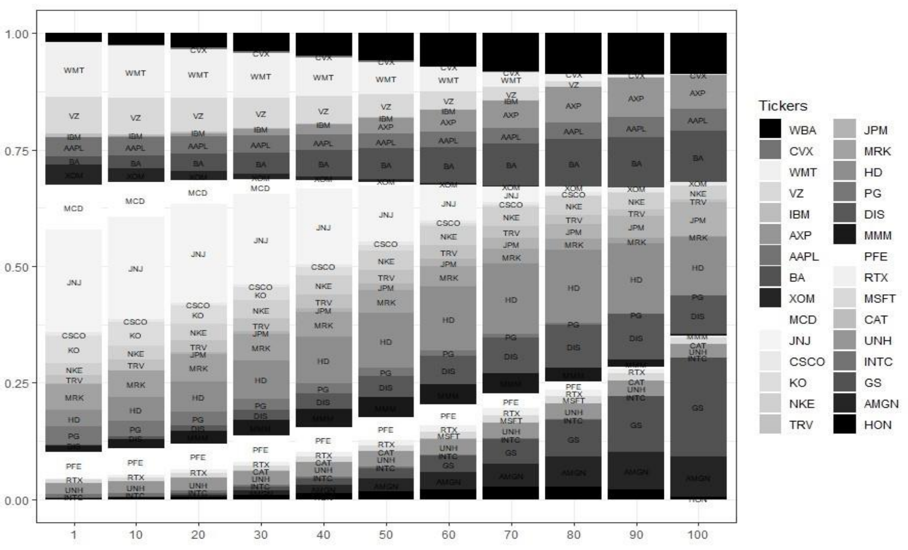

Fig. 2. Distribution of weights within selected portfolios at the efficient frontier quantified using the CVaR model and resampling procedure using GLD distribution as a model of daily returns.

We consider four representative portfolios from each model; for comparison, we have selected four portfolios from each model that share the same expected return to investors. To compare the performance of individual representative portfolios, we quantify average statistics for a period of one year after the investment horizon, namely average return, mean absolute deviation (MAD) and the modified Sharpe ratio from performance measures, where we consider the mean absolute deviation as a measure of risk. Such a statistic is also referred to as the mean absolute deviation ratio. We chose the absolute deviation as a measure of risk (even in the case of the performance measure), as in comparison with the more traditionally used standard deviation and the Sharpe ratio, we are not limited by the assumption of a normal distribution of returns.

Figure 1 and Figure 2 are composite bar graphs where one bar is composed of smaller parts (different shades of gray) which represent the size of the weights of the individual assets in the portfolio. The horizontal axis captures the rank of the individual portfolios, where the first portfolio is the minimum risk portfolio and the last portfolio is the maximum return portfolio. The vertical axis captures the cumulated sum of weights in the portfolio. The assets are displayed in the same order for each bar, the highest in the bar is always the weight of the WBA assets, then the order continues in accordance with the list next to the chart (CVX, WMT, ...) the lowest is always the weight of the HON assets. From the graphical comparison of the structure of selected portfolios, it is clear that the portfolios obtained by resampling procedures are more significantly diversified in comparison with the classical approach. All the more so in the case of portfolios with higher expected returns located in the right half of the individual figures.

Table 20. Quantified average statistics of selected representative portfolios for each investment horizons and tested model.

		One-year investment horizon			Three-year investment horizon			Six-year investment horizon		
Model	**Expected return**	**Mean**	**MAD**	**Sharpe_ Ratio**	**Mean**	**MAD**	**Sharpe _Ratio**	**Mean**	**MAD**	**Sharpe _Ratio**
„Classical“	0,069%	0,001%	0,839%	0,0007	0,073%	0,431%	0,1693	0,044%	0,572%	0,0776
„Classical“	0,083%	0,002%	0,815%	0,0028	0,077%	0,410%	0,1880	0,037%	0,544%	0,0689
„Classical“	0,118%	0,013%	0,884%	0,0150	0,090%	0,401%	0,2256	0,047%	0,661%	0,0716
„Classical“	0,127%	**0,032%**	**0,906%**	**0,0355**	**0,099%**	**0,417%**	**0,2364**	0,068%	0,670%	0,1021
Norm_mult	0,069%	0,003%	0,823%	0,0036	**0,095%**	0,356%	**0,2668**	**0,072%**	0,462%	0,1553
Norm_mult	0,083%	0,004%	0,813%	0,0049	**0,096%**	0,370%	**0,2583**	0,071%	**0,454%**	0,1557
Norm_mult	0,118%	**0,027%**	**0,863%**	**0,0318**	0,094%	0,382%	0,2460	0,082%	0,658%	0,1253
Norm_mult	0,127%	0,027%	0,924%	0,0295	0,087%	0,449%	0,1945	0,098%	0,696%	0,1404
Gld_mult	0,069%	**0,006%**	**0,790%**	**0,0072**	0,087%	**0,356%**	0,2448	0,072%	**0,448%**	**0,1598**
Gld_mult	0,083%	**0,010%**	**0,788%**	**0,0121**	0,091%	**0,365%**	0,2496	**0,072%**	0,455%	**0,1578**
Gld_mult	0,118%	0,026%	0,867%	0,0305	**0,098%**	**0,382%**	**0,2557**	**0,083%**	**0,611%**	**0,1363**
Gld_mult	0,127%	0,023%	0,918%	0,0247	0,094%	0,458%	0,2058	**0,099%**	**0,665%**	**0,1486**

Table 1 shows the quantified average statistics of individual selected portfolios with the same expected return for individual investment horizons. Portfolios with the same expected return for each model are highlighted in the same color. The best values in terms of portfolios with the same expected return in individual investment horizons are highlighted. The data show that portfolios quantified using the resampling procedure achieve better results on average. Only in the case of portfolios with the highest expected return does the "classical" approach achieve better value in the first two investment horizons. Portfolios quantified using the resampling procedure and the GLD distribution are, on average, characterized by lower risk, which may favor such portfolios in the case of longer-term investments.

4 Conclusion

The paper deals with the application of the resampling procedure within the process of portfolio selection in the space of expected return and CVaR using the generation of data from a multivariate random variable. The paper contributes to empirical research by analyzing this procedure on daily data using CVaR model of portfolio selection, which have not been the subject of many contributions so far. The paper also describes a procedure using the generation of data from a multivariate GLD distribution, with most of the empirical research to date dealing mainly with multivariate normal distribution. In this paper we deal with the application of two modifications of such a procedure using a multivariate normal distribution and a multivariate generalized lambda distribution. The performed computational experiments show that the portfolios generated by the resampling procedure are more significantly diversified compared to the "classical" approach. A comparison of the performance of individual portfolios within individual investment horizons shows that, on average, portfolios quantified by the resampling procedure achieve better values compared to the “classical” approach. Portfolios quantified using the GLD distribution have, on average, a lower level of risk. Drawing stronger conclusions will require a more extensive analysis, on the other hand, the obtained results stimulate our interest in further future research in this area.

Acknowledgements

This work was supported by the Grant Agency of Slovak Republic – VEGA grant no. 1/0339/20 „ Hidden Markov Model Utilization in Financial Modeling “.

References

1. Dean, B.: Improved estimation and regression techniques with the generalized lambda distribution. PhD thesis. The University of Newcastle, Australia, Faculty of Science and Information Technology (2013).
2. Delcourt, F. and Petitjean, M.: To What Extent is Resampling Useful in Portfolio Management?. In: Applied Economics Letters, 18, pp. 239-244 (2011).
3. Genz, A., Bretz, F., Miwa, T. and Mi, X.: Package ‘mvtnorm’. [online] [22.05.2021] Dostupné na: <https://cran.r-project.org/web/packages/mvtnorm/mvtnorm.pdf>
4. Lange, A., Sohrmann, C., Jancke, R., Haase, J., Cheng, B., Kovac, U. and Asenov, A.: A general approach for multivariate statistical MOSFET compact modeling preserving

correlations. In: Proceedings of the European Solid-State Device Research Conference 2011, Helsinki: Institute of Electrical and Electronics Engineers, pp. 163-166 (2011).

5. Markowitz, M. H. and Nilufer, U. Resampled Frontiers vs Diffuse Bayes: An Experiment, In: The Journal of Investment Management, 1(4), pp. 9-25 (2003).
6. Michaud, O. R. and Michud, O. R. Efficient Asset Management. Boston, MA: Harvard Business School Press (1998).
7. Michaud, O. R. The Markowitz Optimization Enigma: Is ‘Optimized’ Optimal?. In: SSRN Electronic Journal, 45(1), pp. 31-42. (1989).
8. Pekár, J. and Pčolár, M.: Empirical distribution of daily stock returns of selected developing and emerging markets with application to financial risk management. In: Central European Journal of Operations Research, pp. 1-33 (2021).
9. Quintana, D. Resampled Efficient Frontier Integration for MOEAs. In: Entropy 23(4):422 pp. 1-18 (2021).
10. Rockafellar, R. T. and Uryasev, S.: Optimization of Conditional Value-at-Risk. In: Journal of Risk, 2(3), pp. 21-42 (2000).
11. Scherer, B.: Portfolio Resampling: Review and Critique. In: Financial Analyst Journal, 58, pp. 98-109 (2002).
12. Siswanah, M.: Comparative Analysis of Mean Variance Efficient Frontier and Resampled Efficient Frontier For Optimal Stock Portfolio Formation. IOP Conference Series Materials Science and Engineering, 846:012065, pp. 1-10 (2020).