

Portfolio Optimization Using SPEA2 with Resampling

Sandra García, David Quintana, Inés M. Galván, and Pedro Isasi

Computer Science Department, Carlos III University of Madrid
Avda. Universidad 30, 28911 Leganes, Spain
<http://www.evannai.inf.uc3m.es>

Abstract. The subject of financial portfolio optimization under real-world constraints is a difficult problem that can be tackled using multiobjective evolutionary algorithms. One of the most problematic issues is the dependence of the results on the estimates for a set of parameters, that is, the robustness of solutions. These estimates are often inaccurate and this may result on solutions that, in theory, offered an appropriate risk/return balance and, in practice, resulted being very poor. In this paper we suggest that using a resampling mechanism may filter out the most unstable. We test this idea on real data using SPEA2 as optimization algorithm and the results show that the use of resampling increases significantly the reliability of the resulting portfolios.

1 Introduction

The problem of choosing the right combination of financial assets has been the subject of research for a long time and it is one of the most active research lines in finance. This is often framed as a multiobjective optimization problem where the investor tries to find the right set of portfolios with the best risk/return profiles.

A large portion of the academic literature on this subject builds on the seminal work by Markowitz [5, 6]. The approach suggested by this author works under some assumptions that allow the problem to be tackled with quadratic programming. Unfortunately, these assumptions do not hold in the real-world, which calls for alternatives. That is the reason why the framework of evolutionary multiobjective optimization is getting traction on this area [1, 2, 13]. In all these references, the authors use evolutionary multiobjective algorithms to evolve sets of solutions that minimize risk and maximize return. The first one, introduces a customized hybrid version NSGA-II and the last two compare the performance of different multiobjective algorithms.

One of the most important factors that asset managers face when they have to assess the results provided by any of the above-mentioned methods is stability. Very often, the expected efficient frontier lies far from the actual one as the forecasted risk/return profile of the portfolios is not accurate. This problem is one the major reasons why some practitioners mistrust the mentioned approach and the search for solutions has cleared the way for the field of robust portfolio optimization. The main contribution of this paper is the introduction of a new resampling mechanism that reduces the risk mentioned.

When we forecast the risk and return of a specific portfolio, we rely on estimates for the expected returns of individual assets and the variance-covariance matrix. The

forecasts, which are likely to be inaccurate, are usually based on past data which may not be representative to predict the future due to, for instance, the presence of outliers.

In this scenario, there are several potential ways to approach the problem. The main two are either putting an emphasis on having robust estimates for the above mentioned parameters [7] or implementing a system that deals with uncertainty in the estimation process [8, 9]. The approach suggested in this paper is a new technique that falls in the latter category. We will use an evolutionary multiobjective algorithm enhanced with a resampling mechanism that changes the parameters of the fitness function during the evolution process with the aim to obtain robust solutions. We consider that using a multiobjective genetic algorithm that exposes candidate solutions to different scenarios will improve the reliability of the resulting Pareto front. The use of resampling in the context of portfolio optimization is not new [11, 12]. The most comparable approach is described by Idzorek [4] who suggests using combining traditional quadratic programming (QP) with Monte Carlo simulation to derive a set of fronts that are subsequently merged into a single solution. We understand that resampling within the context of multiobjective evolutionary algorithms is a better strategy as it would allow real-world constraints intractable by QP while, at the same time, approximating the efficient frontier in single run.

The rationale for the approach is that optimizing for a single scenario bears the risk of getting solutions that are hyper-specialized and might be extremely sensitive to deviations in the parameters used in the fitness evaluation. Given that it is almost certain that we will not be able to predict accurately the behaviour of all the assets, we could consider the alternative of targeting portfolios that offer appropriate risk/return tradeoffs under different scenarios. Even if the solutions do not seem to be as good as the specialized ones under the expected scenario, they might be more reliable in practice. Once we replace the expected parameters of the models with the observed ones, the actual risk and returns might be more likely to be closer to the expected values, hence providing more value to the decision maker.

The choice of performance metrics is a very important issue in multiobjective optimization [15]. It is generally admitted that there is no single metric that can be used to evaluate different objectives simultaneously. In addition to that, we also face the lack of standard robustness metrics that we could use in this context. For this reason, we introduce a new one that accounts for the average difference on the objective space using the estimates for the parameters vs. using their real values, across all the solutions.

The rest of the paper is organized as follows. First, we make a formal introduction to the portfolio optimization problem. Then, we describe in detail our approach and the proposed metric to evaluate the robustness of the solutions. That will be followed by the experimental results and a section devoted to summary and conclusions.

2 Portfolio Optimization's Problem Definition

The Modern Portfolio Theory was originated in the article published by Harry M. Markowitz, in 1952 [5]. In general, the portfolio optimization problem is the choice of an optimum set of assets to include in the portfolio and the distribution of investor's wealth among them. Markowitz [6, 10] assumed that solving the problem requires the

simultaneous satisfaction of maximizing the expected portfolio return $E(R_p)$ and minimizing the portfolio risk (variance) σ_p^2 . This can be formally defined as:

- Minimize the **risk** (variance) of the portfolio:

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \sigma_{ij} \quad (1)$$

- Maximize the **return** of the portfolio:

$$E(R_p) = \sum_{i=1}^n w_i \mu_i \quad (2)$$

- Subject to:

$$\sum_{i=1}^n w_i = 1 \text{ and } 0 \leq w_i \leq 1; i = 1 \dots n \quad (3)$$

where n is the number of available assets, μ_i the expected return of the asset i , σ_{ij} the covariance between asset i and j , and w_i are the weight giving the composition of the portfolio. The constraints referenced in equations 3 require the full investment of funds and prevents the investor from shorting any asset.

3 Parameter Resampling Strategy

In this work, the problem of obtaining stable and robust portfolios is tackled using an evolutionary multiobjective algorithm, specifically, SPEA2 [14]. This choice is justified by the results reported in [1] and [1, 13]. Both compare several algorithms in the domain and conclude that SPEA2 offers the best results.

The encoding chosen will represent each portfolio as a vector of real numbers, this means that SPEA2 will work with real elements instead of binary ones. Each of these numbers represents the percentage of investment per asset (also called weight: w_i where $i = 1 \dots n$ and n is the number of investable assets). Here, each portfolio will be represented by a single element of the population. Every individual must meet the constraints from eq. 3 and therefore, will be repaired after applying the genetic operators.

As we mentioned, the main contribution of this paper is the way that we handle the fitness function. The starting point for the evaluation of portfolios is the framework introduced in section 2, where the fitness of each individual is determined by evaluating the two objective functions: return $E(R_p)$ and risk σ_p .

As it is apparent from equations 1 and 2, these functions are very dependent on the values of the expected asset returns and the variance-covariance matrix. One of the most important challenges in portfolio context is the dependence of the solutions on the estimates for these parameters. Given the difficulty inherent to financial forecasting, it is unlikely that these parameters are accurate. This lack of accuracy is likely to result in a set of portfolios that end up behaving in an unexpected way. Fig. 1 shows a real example of one Pareto front where the solutions are evaluated using the forecasted parameters (in red) and the real parameters (in green). For this reason, we suggest altering the fitness function in a way that it discards those portfolios that, under normal circumstances, could potentially show bad performance.

The basic idea behind the solution that we suggest is to keep changing the parameters of the fitness functions during the evolution process. If these parameters take values

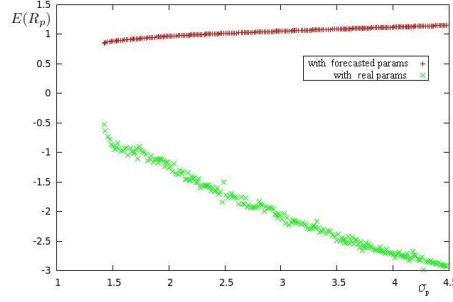


Fig. 1. Solution evaluated with forecasted and real parameters

that are likely to be close to the real ones, evolution would tend to favour those individuals that tend to have a good performance in all circumstances (they will not be necessarily the best for the real value of the parameters, that will only be known a posteriori). Those solutions that tend to be good most of the time but are particularly bad in some scenarios may have to face them and therefore, would be weeded out of the population with high probability due to their low fitness. A key factor for the success of this strategy is the mechanism used to generate the scenarios.

The approach that we use to generate these scenarios is non-parametric bootstrap. Our algorithm used a time window to resample data, instead of using all the data to derive a single estimate for the parameters. The resampling process selects a random set of time periods that has the same size as the original window (each period might be selected more than once). Then, we average the returns for those time periods and compute the variance-covariance matrix. Whenever we do this, we generate new estimates for the parameters that are based on past data. These estimates can subsequently be used to calculate the risk and return of the portfolios. The process generates a new sample set S' with size N'_s from the original sample S with size N_s . While $N'_s \neq N_s$ one instance X_i is selected randomly from S and added to the new set $S' = S' + X_i$. At the beginning, $S' = \emptyset$ and $N'_s = 0$.

Hence, when we use resampling we create a set of "likely scenarios" that prevents the solution from hyper-specializing in just one, making the solution more robust.

4 Evaluation Metric

In the context studied in this work, the metrics most commonly used, as Hypervolume (HV) and Spread, can not be used to measure the quality of solutions from a robustness point of view. HV is an indicator of the size of the dominated space and measures the volume enclosed by the union of the points in the front; the bigger value it gives, the better HV it has. The Spread is a diversity indicator which measures the extent of spread achieved among the obtained solutions so, the lower value, the more evenly distributed is the front.

In order to measure the robustness of the solutions, we have defined a new metric, named **Estimation Error (EE)**. The aim is to evaluate the average difference between

the expected risk and return for every portfolio in the efficient frontier and the actual risk and return a posteriori once the real value of the parameters is observed. That is, the difference between the estimates for t_n based on data from t_1 to t_{n-1} , and the actual values at t_n . This metric is calculated measuring the average Mahalanobis distance (d_M) between the forecasted pair $(E(R_p), \sigma_p^2)$ for each portfolio, named $\bar{x}_i$, and the $(E(R_p), \sigma_p^2)$ for the same portfolio computed with the real parameters, named $\bar{x}_i'$ (the observed in t_n). The lower the value is, the higher is the reliability of the original front. Formally, it can be expressed as follows:

$$EE = \frac{\sum_{i=1}^p [d_M(\bar{x}_i, \bar{x}_i')]^2}{p} \quad (4)$$

where p is the number of portfolios in the Pareto front.

5 Experimental Results

In this section we test the aforementioned approach on a specific asset allocation problem. The sample consists of 320 monthly returns for eight broad financial indexes representing that many asset classes. The series of monthly returns cover from May, 1980 to December, 2006 and the source is *DataStream*. The list of indexes used in this work is the following: Frank Russell 2000 Growth (FRUS2GR), Frank Russell 1000 Value (FRUS1VA), Frank Russell 1000 Growth (FRUS1GR), S&P GSCI Commodity Total Return (GSCITOT), MSCI EAFE (MSEAFEL), BOFA ML CORP MSTR (\$) (MLCORPM) and, BOFA ML US TRSY /AGCY MSTRAAA (\$) (MLUSALM).

As we have mentioned, SPEA2 has been used for experiments. The code was developed in Java under jMetal [3], that includes SPEA2. Simulated Binary Crossover and Polynomial Mutation are used as crossover and mutation operators. Probabilities are respectively fixed to 0.125 and 0.9.

The SPEA2+resampling strategy is compared to SPEA2. For both strategies, resampling and no-resampling, experiments varying the population size (50, 100 and, 200) and the number of evaluations (5000, 10000 and 20000 respectively) have been carried out. Experiments with and without resampling are, respectively, called as R-50, R-100, R-200, and NR-50, NR-100, NR-200 where the numbers 50, 100 and 200 denote the population size.

A sliding window must be used when fitness function is evaluated using the resampling technique. Each window has a size n of 200 monthly returns. The sets t_1 to t_{n-1} are used in the algorithm to get the final solutions and the period t_n is used to evaluate them. In our experimentation the window is moved 120 times (one month at a time) taking the interval from 30/05/1980 to 29/12/2006. The algorithm is run 30 times per window getting, for each window, a 30 solution set.

We report in table 1 the average results for the metrics referenced in 4. It shows the average, median, variance, maximum and minimum values of the metrics to compare experiments with and without resampling. To calculate HV and Spread, the extremes of optimal front are approximated with the limit points from the whole set of the experiments.

Table 1. Metrics values

EE	R-50	NR-50	R-100	NR-100	R-200	NR-200
Average	3.58389	3.62739	3.58586	3.75097	3.61118	3.76668
Median	3.75999	3.84983	3.77780	3.92876	3.80308	3.94310
Variance	0.18572	0.22016	0.19541	0.16160	0.19780	0.15766
Maximum value	3.94303	3.95819	3.96495	3.97959	3.97590	3.98965
Minimum Value	1.83713	1.84397	1.80752	2.03810	1.82528	2.07942
HV	R-50	NR-50	R-100	NR-100	R-200	NR-200
Average	0.51551	0.50941	0.47368	0.50176	0.47223	0.49715
Median	0.53966	0.45816	0.42839	0.45258	0.42656	0.44895
Variance	0.01767	0.01728	0.01700	0.01645	0.01679	0.01616
Maximum value	0.77764	0.78560	0.74998	0.76184	0.74420	0.75424
Minimum Value	0.22626	0.30859	0.24046	0.30668	0.23544	0.30320
SPREAD	R-50	NR-50	R-100	NR-100	R-200	NR-200
Average	0.89554	0.80186	0.88501	0.47855	0.88513	0.46028
Median	0.89228	0.79946	0.88210	0.49075	0.87906	0.47166
Variance	0.00715	0.01020	0.00997	0.00909	0.01034	0.00951
Maximum value	1.22227	1.11792	1.33979	0.64023	1.43464	0.61230
Minimum Value	0.57814	0.49306	0.59058	0.25040	0.52798	0.22788

We notice that all the experiments with resampling present smaller EE averages and variances. This means that resampled solutions behave as expected to a greater degree than non-resampled ones. In terms of HV, the averages for "NR-100" and "NR-200" are a bit larger than the ones observed for "R-100" and "R-200" but the HV of "R-50" and "NR-50" is similar. The average Spread for the SPEA2+resampling is larger than the equivalent metric for non-resampled solutions, which means that the latter tends to have a better distribution. We also see that EE grows with the population size. Regardless of this size, the resampling method always provides lower EE and more Spread than the basic one and these differences are significant at 1% using Wilcoxon test. This is due to the emphasis of resampling technique on the selection of stable individuals, which might be found in specific sections of the front, while the basic approach is not limited by this factor. The same test shows that the HV of "NR-100" and "NR-200" are higher, with a significance difference of 1%, than the ones of "R-100" and "R-200", however this difference does not exist between "R-50" and "NR-50". So, even if the resampled solutions do not seem to be as good as the others in terms of HV and spread, the EE metric shows that they might be more robust and reliable in practice providing more value to the decision maker.

In order to illustrate the shape of the fronts, we show two fronts example in fig. 2, one from "R-50" (in green) and the other one from "NR-50" (in red).

The results presented above have shown that resampling always improves significantly the reliability of the solutions as measured by average error committed predicting the performance of the final set of portfolios. However, this robustness comes at a cost

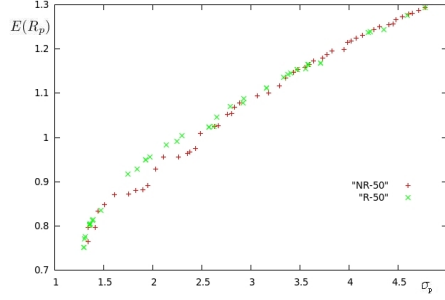


Fig. 2. Comparison between resampled and non-resampled fronts

as non-resampled portfolios tend to be distributed more evenly along the front and for some cases the hypervolume tends to be higher.

6 Conclusions

Portfolio optimization is one of the most active research lines in finance and the choice of the right combination of financial assets can be framed as a multiobjective optimization problem.

Portfolio managers often face the problem that the expected efficient frontier derived from their forecasts for future returns is subject to uncertainty. This means that if the real parameters differ from the forecasted ones, the risk and returns the portfolios included in the estimated efficient frontier might deviate substantially from the predictions. This could result extreme underperformance of some portfolios. This uncertainty is one of the major reasons why some practitioners mistrust quantitative methods based on modern portfolio theory, and the search for solutions has cleared the way for the development of robust portfolio optimization.

We presented a new resampling mechanism based on non-parametric bootstrap that, used in combination with MOEAs such as SPEA2 should lower the likelihood of getting unstable or not robust solutions. The mechanism resamples past data to generate different scenarios that are used during the evolutionary process to evaluate the portfolios in the population. Those portfolios that perform very poorly in any of these scenarios get discarded by the algorithm. This results in a final front consists of individuals that tend to offer good performance in any circumstance.

The approach was tested on real data on a sample of monthly returns for eight indexes representing different broad investment categories including stock, bonds, etc. We compared the performance of SPEA2 vs SPEA2+Resampling. The results show that resampling enhances significantly the reliability of the solutions as measured by average error committed predicting the performance of the final set of portfolios. This is very important as it indicates decision makers could rely more in the robust front to pick the right portfolio according to their preferences.

Even though these results are encouraging, there are several issues left open that could lead to future extensions of this work. Among them, the performance of the re-

sampling mechanism using other MOEAs, the scalability of the results with the size of the number of investment alternatives, or the performance of the approach once the decision maker faces cardinality restrictions to the amount he can invest in a specific asset class.

Acknowledgements

The authors acknowledge financial support granted by the Spanish Ministry of Science under contract TIN2008-06491-C04-03 (MSTAR) and Comunidad de Madrid (CCG10-UC3M/TIC-5029).

References

1. K.P. Anagnostopoulos and G. Mamanis. The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, In Press, Corrected Proof, 2011.
2. K. Deb, R. E. Steuer, R. Tewari, and R. Tewari. Bi-objective portfolio optimization using a customized hybrid nsga-ii procedure. In *EMO*, pages 358–373, 2011.
3. J.J. Durillo, A.J. Nebro, and E. Alba. The jmetal framework for multi-objective optimization: Design and architecture. In *CEC 2010*, pages 4138–4325, Barcelona, Spain, July 2010.
4. Thomas M. Idzorek. Developing robust asset allocations. 2006.
5. H.M. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
6. H.M. Markowitz. *Portfolio Selection: efficient diversification of investments*. John Wiley & Son, 1959.
7. Cdric Perret-Gentil and Maria-Pia Victoria-Feser. Robust mean-variance portfolio selection. Fame research paper series, International Center for Financial Asset Management and Engineering, April 2005.
8. Georg Pflug and David Wozabal. Ambiguity in portfolio selection. *Quantitative Finance*, 7(4):435–442, 2007.
9. H. Tütüncü Reha and M. Koenig. Robust asset allocation. *Annals OR*, 132(1-4):157–187, 2004.
10. F.K. Reilly and C. K. Brown. *Investment Analysis and Portfolio Management*. The Dryden Press, 1997.
11. David Ruppert. Statistics and finance: An introduction. *Journal of the American Statistical Association*, 101:849–850, June 2006.
12. H. Shiraishi. Resampling procedure to construct value at risk efficient portfolios for stationary returns of assets. 2008.
13. P. Skolpadungket, K. Dahal, and N. Harnpornchai. Portfolio optimization using multi-objective genetic algorithms. In *Proceeding of 2007 IEEE Congress on Evolutionary Computation*, pages 516–523, 2007.
14. E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the strength pareto evolutionary algorithm. Technical report, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, 2001.
15. E. Zitzler and L. Thiele. An evolutionary algorithm for multiobjective optimization: The strength pareto approach. Technical Report 43, Gloriastrasse 35, CH-8092 Zurich, Switzerland, 1998.