A Comparative Study of Multi-Objective Evolutionary Algorithms to Optimize the Selection of Investment Portfolios with Cardinality Constraints

Feijoo Colomine Duran[1], Carlos Cotta[2], and Antonio J. Fernández-Leiva[2]

[1] Universidad Nacional Experimental del Táchira (UNET)
Laboratorio de Computación de Alto Rendimiento (LCAR), San Cristóbal, Venezuela
fcolomin@unet.edu.ve

[2] Dept. Lenguajes y Ciencias de la Computación, ETSI Informática,
University of Málaga, Campus de Teatinos, 29071 - Málaga, Spain.
{ccottap,afdez}@lcc.uma.es

Abstract. We consider the problem of selecting investment components according to two partially opposed measures: the portfolio performance and its risk. We approach this within Markowitz's model, considering the case of mutual funds market in Europe until July 2010. Comparisons were made on three multi-objective evolutionary algorithms, namely NSGA-II, SPEA2 and IBEA. Two well-known performance measures are considered for this purpose: hypervolume and R_2 indicator. The comparative analysis also includes an assessment of the financial efficiency of the investment portfolio selected according to Sharpe's index, which is a measure of performance/risk. The experimental results hint at the superiority of the indicator-based evolutionary algorithm.

1 Introduction

There are several theoretical studies related to risk-return interaction. The potential loss of performance or investment is not static, but it always depends on market developments. In the literature, we can find several proposals that model this scenario. For example, Markowitz's model [1] has become a key theoretical framework for the selection of investment portfolios. However, its application in practice has not been as extensive, mainly due to the mathematical complexity of the method.

Markowitz's model with multiple objectives is expressed as:

$$\min \sigma^2(R_p) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}, \quad \max E(R_p) = \sum_{i=1}^{n} w_i E(R_i) \tag{1}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1, \text{ and } w_i \geqslant 0 \; (i = 1, \cdots, n) \tag{2}$$

where w_i is the investor's share of the budget for the financial asset i (to be found), $\sigma^2(R_p)$ is the variance of the portfolio p, and σ_{ij} is the covariance between the returns of the values i and j. $E(R_p)$ is the expected return or portfolio p, which are the set of proportions that minimize the risk of the portfolio and and its corresponding value. The set of pairs $[E(R_p), \sigma^2(R_p)]$ or combinations of all risk-return efficient portfolios is called the efficient frontier. Once known, the investor chooses according to his preferences the optimal portfolio. In this model some other restrictions can be added, such as cardinality (a maximum of K is non-zero weights) or limits on the percentage of an asset allocation. In this paper we focus on the constraints of the first type:

$$\sum_{w_i>0} 1 = K \tag{3}$$

The research considers a comparative study between different multi-objective evolutionary algorithms. More specifically the comparative study considers three algorithms: SPEA2 [2], NSGA-II [3], and IBEA [4]. We also address the selection of a point in the Pareto front using Sharpe's index [5] whose expression is:

$$S_p = \frac{E(R_p) - R_0}{\sigma_p} \tag{4}$$

This index is a risk-return ratio. The numerator is the excess return defined by the difference between the yield on the portfolio ($E(R_p)$) and the risk free rate (R_0) in the same period of assessment. The portfolio risk is measured by the standard deviation of this (σ_p). That is, indicates the yield premium offered by a portfolio per unit of total risk of the same. It follows that the higher the risk-reward ratio, the greater the success of fund management.

The objective of this goal is to assess comparatively the performance of several multi-objective EAs on this problem scenario. For this purpose, we will firstly overview some related work in next section.

2 Related work

There is a plethora of works in the literature dealing with the use of MOEAs in the area of investment portfolio optimization. Without being exhaustive, we can firstly cite the work by Diosan [6], who makes a comparison between PESA [7], NSGA II and SPEA2, and through empirical results suggests the adequateness of PESA. See also [8] for a comparison of these techniques in this context. Chang et al. [9] compare the use of different meta-heuristics (both evolutionary and local-search ones) for finding the efficient frontier by adding cardinality constraints. Perez et al. [10] make in turn a comparison between MOGA [11], NPGA [12] and NGGA [13] concluding the superiority of NGGA. On the other hand, Doerner et al. [14] make a comparison between Pareto Optimization Using Ant Colony (PACO), simulated annealing and NSGA [15] in terms of both quality and computational cost, suggesting the superiority of PACO. Ehrgott et al. [16] propose an interesting variation of Markowitz's model adding additional objective functions to consider the individual preferences of the investor, and find genetic algorithms to perform better than local-search techniques. Skolpadungket et al. [17] perform a comparison between VEGA [18], MOGA, NSGA II and SPEA2 in the context of financial portfolio optimization, and find that SPEA2 provide the best performance.

3 Material and Methods

In the following we will address the data and algorithms used in the experimentation, as well as the performance measures considered for evaluating performance.

3.1 Data Analyzed

We consider data corresponding to mutual funds in Europe. More precisely, these data comprise the stock value of the funds, sampled on a monthly basis for five years. This period of time is long enough to span a full market cycle, with rises and falls, but not large enough to comprise profound changes in the reality of each of the fund shares, thus making past information little representative for foreseeing future performance. Based on this analysis we take a sample at the discretion of choosing funds that are no older than five years and are still available on the market. Table 1 shows the various funds used. The risk-free return R_0 is referenced is 0.04.

Table 1. Mutual funds considered [19]

	Mutual Fund Europe
1	DFA United Kingdom Small Compan
2	Eastern European Equity A (VEEE)
3	Eastern European Equity C (VEEC)
4	Henderson European Focus A (HFE)
5	Henderson European Focus B (HFE)
6	ING Russia A (LETRX)
7	JPMorgan Russia A (JRUAX)
8	JPMorgan Russia Select (JRUSX)
9	Metzler Payden European Emergin
10	Mutual European A (TEMIX)
11	Mutual European B (TEUBX)
12	Mutual European C (TEURX)
13	Mutual European Z (MEURX)
14	Royce European Smaller Companie
15	Third Millennium Russia I (TMRI)

3.2 Algorithmic Methods

The optimization of investment portfolios according to a variety of performance-risk profiles lends itself very well to multi-objective optimization techniques in general, and multi-objective evolutionary algorithms (MOEAs). Aiming to compare the performance of three different MOEAs on the same problem setting and using the same experimental data, we have considered the following techniques:

1. Second-generation Pareto-based MOEAs: NSGA-II (Non-dominated Sorting Genetic Algorithm II) [3] and SPEA2 (Strength Pareto Evolutionary Algorithm 2) [2]. These MOEAs are based on the notion of Pareto-dominance, used for determining the solutions that will breed and/or the solutions that will be replaced. furthermore, as second-generation techniques they exploit elitism (an external archive of non-dominated solutions in the case of SPEA2, and a plus-replacement strategy in the case of NSGA-II). More precisely, NSGA-II sorts the population in non-domination levels (performing binary tournament on the so-obtained ranks for selection purposes), and uses crowding for performing replacement and spreading the Pareto front. As to SPEA2, it features an external archive of solutions which is used to calculate the "strength" of each individual i (the number of solutions dominated by or equal to i, divided by the population size plus one). This is used for selection purposes (aiming to minimize the strength of solutions which are non-dominated by tentative parents). SPEA2 also includes a nearest-neighbor density estimation technique to spread the front, and a sophisticated archive update strategy to preserve boundary conditions.
2. Indicator-based MOEAs: IBEA (Indicator-Based Evolutionary Algorithm) [4] attempts to incorporate practical decision-making and privileged information when searching for Pareto solutions. The question that arises is how to concentrate the search in regions of the Pareto front that are of interest to the person responsible for taking the decision. This is done in IBEA by maximizing (or minimizing) some performance indicator. By doing so, IBEA can be considered a collective approach, where selective pressure is exerted to maximize the performance of the whole population. In this work, we have considered an IBEA based on the ε-indicator [20].

As to the parameters considered, these are described in Table 2.

3.3 Performance Indicators

The evaluation of the performance of a MOEA is in itself a multi-objective problem, that can be approached in multiple ways. We have considered two well-known performance indicators: the hypervolume indicator [21] and the R_2 indicator [22]. The first one provides an indication of the region in the fitness space that is dominated by the front, and which must be maximized for better performance. As to the second indicator, it estimates the extent to which a certain front approximates another one (the Pareto-optimal front or a reference

Table 2. Parameterization considered in the experimentation

Parameter	Value
representation	binary
number of genes	15
gene size	10 bits
size of chromosome	150
population size	300
generations	100
selection type	tournament/elitist
crossover operator	2-Point Crossover
crossover probability	0.8
mutation operator	bitflip
mutation probability	0.0666

front if the former is unknown). We have considered the unary version of this indicator, taking the combined NSGA-II/SPEA2/IBEA Pareto front as a reference set. Being a measure of distance to the reference set, R_2 must be minimized for better performance

4 Experimental Results

Experimentation has been done before the three algorithms described, namely, NSGA-II, SPEA2 and IBEA, using the PISA library [23], the parameters described in Table 2, and the data described in Table 1. Four different values of the cardinality constraint K have been considered, namely $K \in \{2, 4, 7, 15\}$. For each algorithm and value of K, thirty runs have been done.

4.1 Analysis of the Pareto Front

The first part of the experimentation focused on the analysis of the obtained Pareto front. The three algorithms behave very similarly, with slight differences is the high-risk end of the fronts. To analyze in more detail these results we have applied the two performance indicators mentioned in Sect. 3.3, namely hypervolume and R_2. In the first case we have used as reference a point of maximum risk and minimum benefit. As to the latter, distance is measured against the best-known front (the combined front front the three algorithms). Figures 1–2 shows the distribution of values of the indicators for each algorithm for the two extreme cardinality values ($K = 2$ and $K = 15$).

Inspection of these figures indicates that SPEA2 provides better performance for a low value of K. In this constrained scenario, SPEA2 provides a broader dominance of the performance-risk fitness space, and is globally closer to the best-known Pareto front. However, in the other end of the spectrum (high value of K, and hence less constrained portfolios) IBEA stands out as a more effective

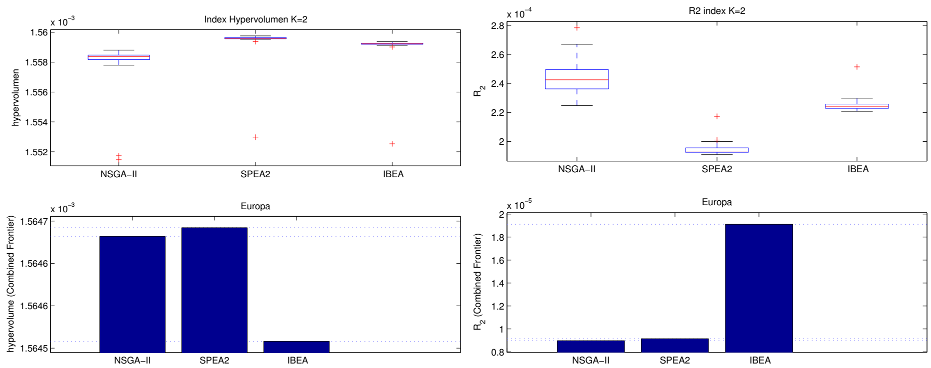

Fig. 1. Hypervolume and R_2 indicator for $K = 2$

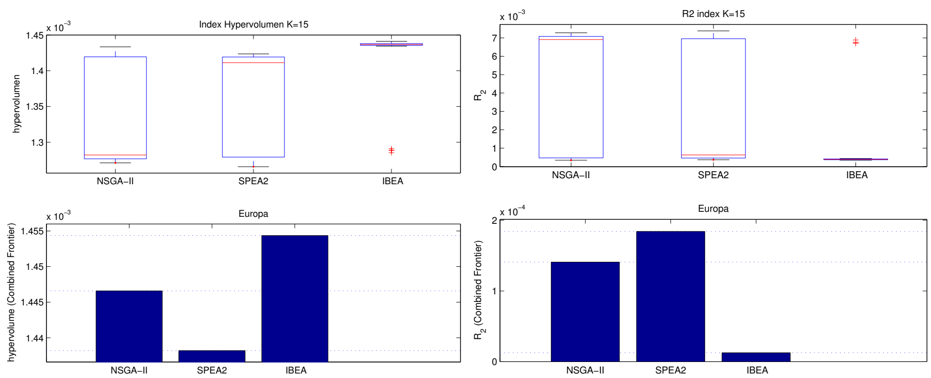

Fig. 2. Hypervolume and R_2 indicator for $K = 15$

algorithm. This is also the case for the intermediate values $K = 4$ and $K = 7$. In all cases, these performance differences can be shown to be statistically significant (at $\alpha = 0.05$) via the use of a non-parametric Wilcoxon test [24].

4.2 Analysis through Sharpe's ratio

Sharpe's index is used for decision-making once a front is obtained; it allows picking a single solution out of the whole efficient front, trying to maximize excess profit per risk unit. Geometrically, this can be interpreted as finding the straight line with highest slope that is tangent to the front and passes through the risk-free point $(0, R_0)$. This is obviously influence by the shape of the front; the analysis of the results through this index can thus be used how efficient are the MOEAs in terms of the performance-risk profile arising from this decision-making procedure. Figure 3 shows the distribution of Sharpe's index values for the different cardinality values.

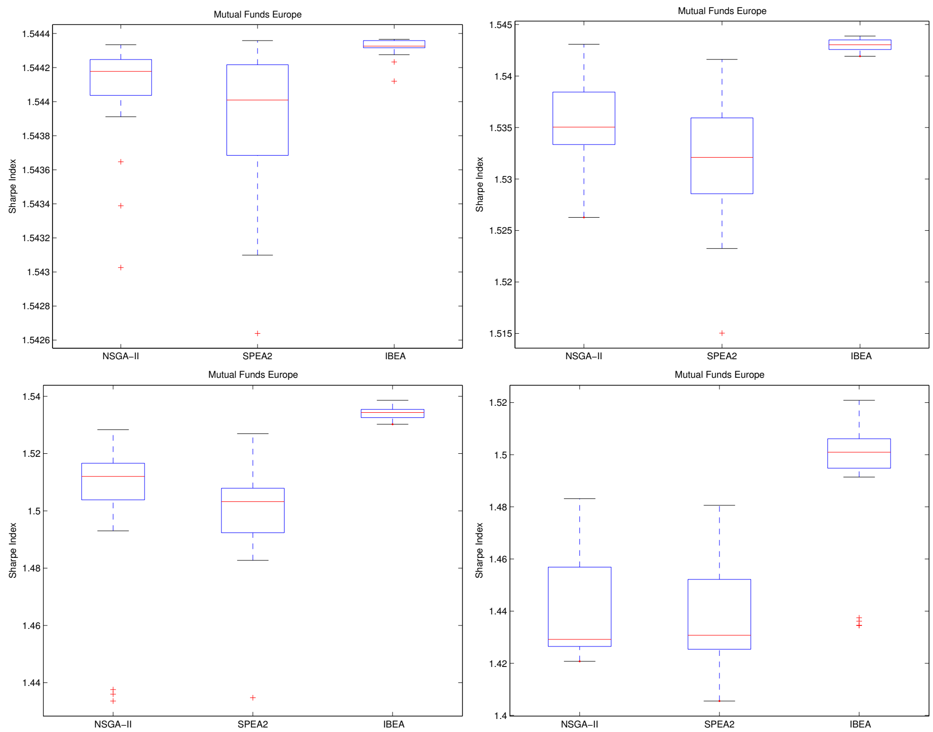

Fig. 3. Sharpe's index distribution for MOEAs under Markowitz's model. Top row: $K = 2$ (left) and $K = 4$ (right). Bottom row: $K = 7$ (left) and $K = 15$ (right).

Visual inspection of these results indicate a substantial advantage for IBEA. This advantage is again shown to be statistically significant via the use of a Wilcoxon ranksum test ($\alpha = 0.05$). We can also make an analysis of how different algorithms behave when varying the risk-free return component (R_0). In this case, it can be observed that as the cardinality rises, IBEA returns slightly better results for increasing risk-free values.

5 Conclusions

The problem of portfolio optimization is a natural scenario for the use of multiobjective evolutionary algorithms, in which their power and flexibility can be readily exploited. In this sense this paper has analyzed three different state-of-the-art MOEAs, namely NSGA II, SPEA2 and IBEA under a common experimental framework centered in mutual funds in Europe. While the three algorithms provide a variety of solution profiles that can be considered optimal in a Pareto sense, a performance analysis conducted under two specific indicators (hypervolume and R_2) indicate that IBEA performs significantly better, in particular when the cardinality constraint K does not take an extremely low value. This can be interpreted in terms of the exploration capabilities of the multi-objective optimizer for the richer (less-constrained) fitness landscapes. As an additional means of comparison among the MOEAs, we have considered the outcome of a decision-making process based on the use of Sharpe's index. Again, IBEA stands out, indicating that it provides a better exploration capability in the area of fitness space around the knee of the front.

In future work we intend to explore other variants of the optimization problem by adding, e.g., maximum and minimum rates of investment. We also intend to study other model variations such as Jensen's alpha [25] under a market model such as CAPM (Capital Asset Pricing Model [26]).

Acknowledgements This work is partially supported by Spanish MICINN under project ANYSELF (TIN2011-28627-C04-01), and by Junta de Andalucía under project TIC-6083.

References

1. Markowitz, H.M.: Portfolio selection. Journal of Finance **7** (1952) 77–91
2. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm. In Giannakoglou, K., et al., eds.: EUROGEN 2001. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Prob- lems, Athens, Greece (2002) 95–100
3. Deb, K., Agrawal, S., Pratab, A., Meyarivan, T.: A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II. In Schoenauer, M., et al., eds.: 6th Parallel Problem Solving from Nature (PPSN VI). Volume 1917 of Lecture Notes in Computer Science., Paris, France, Springer (2000) 849–858
4. Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In Yao, X., et al., eds.: 8th Parallel Problem Solving from Nature (PPSN VIII). Volume 3242 of Lecture Notes in Computer Science., Birminghan, UK, Springer (2004) 832–842
5. Sharpe, W.F.: Mutual fund performance. Journal of Business **39** (1966) 119–138
6. Diosan, L.: A multi-objective evolutionary approach to the portfolio optimization problem. In: Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Vol-2 (CIMCA-IAWTIC'06) - Volume 02, Washington, DC, USA, IEEE Computer Society (2005) 183–187
7. Corne, D., Knowles, J., Oates, M.: The pareto envelope based selection algorithm for multiobjective optimization. In Schoenauer, M., et al., eds.: 6th Parallel Problem Solving from Nature (PPSN VI). Volume 1917 of Lecture Notes in Computer Science., Paris, France (2000)
8. Anagnostopoulos, K.P., Mamanis, G.: A portfolio optimization model with three objectives and discrete variables. Comput. Oper. Res. **37** (July 2010) 1285–1297
9. Chang, T.J., Meade, N., Beasley, J., Sharaiha, Y.: Heuristics for cardinality constrained portfolio optimisation. Computers and Operations Research **27** (2000) 1271–1302

10. Pérez, M.E., del Olmo, R., Herrera., F.: The formation of efficient portfolios using multiobjective genetic algorithms. Technical Report DECSAI-00-01-26, Dept. CCIA, Universidad de Granada (2000)
11. Fonseca, C.M., Fleming, P.J.: Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In Forrest, S., ed.: Genetic Algorithms: Proceedings of the Fifth International Conference, San Mateo, California, University of Illinois at Urbana-Champaign, Morgan Kauffman Publishers (1993) 416–423
12. Horn, J., Nafpliotis, N., Goldberg, D.E.: A niched pareto genetic algorithm for multiobjective optimization. In: First IEEE Conference on Evolutionary Computation, IEEE (1994) 82–87
13. Rendn, M.V., Charre., E.U.: A non-generational genetic algorithm for multiobjective optimization. In Bck, T., ed.: Proceedings of the Seventh International Conference on Genetic Algorithms, San Mateo, California, Michigan State University, Morgan Kaufmann (1997) 658–665
14. Doerner, K., Gutjahr, W., Hartl, R., Strauss, C., Stummer, C.: Pareto ant colony optimization: A metaheuristic approach to multiobjective portfolio selection. Annals of Operations Research **131** (2004) 79–99
15. Srinivas, N., Deb, K.: Multiobjective optimization using nondominated sorting in genetic algorithms. Evolutionary Computation **2** (1994) 221–248
16. Ehrgott, M., Klamroth, K., Schwehm, C.: An MCDM approach to portfolio optimization. European Journal of Operational Research **155**(3) (June 2004) 752–770
17. Skolpadungket, P., Keshav, D., Napat, H.: Portfolio optimization using multiobjective genetic algorithms. IEEE Congress on Evolutionary Computation (2007) 516–523
18. Schaffer, J.D.: Multiple objective optimization with vector evaluated genetica lgorithms. in genetic algorithms and their applications. In Grefenstette, J., ed.: Proceedings of the First International Conference on Genetic Algorithms, Hillsdale NJ, Lawrence Erlbaum (1985) 93–100
19. Yahoo Finance: Mutual funds center. 2011
20. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., Grunert da Fonseca, V.: Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE Transactions on Evolutionary Computation **7**(2) (2003) 117–132
21. Zitzler, E., Thiele, L.: Multiobjective optimization using evolutionary algorithms. a comparative case study. In Eiben, A.E., et al., eds.: 5th Parallel Problem Solving from Nature (PPSN V). Volume 1498 of Lecture Notes in Computer Science., Amsterdam, Springer (September 1998) 292–301
22. Hansen, M., Jaszkiewicz, A.: Evaluating the quality of approximations to the nondominated set. Technical Report IMM-REP-1998-7, Institute of Mathematical Modelling Technical University of Denmark (1998)
23. Bleuler, S., Laumanns, M., Thiele, L., Zitzler, E.: PISA—A Platform and Programming Language Independent Interface for Search Algorithms. In Fonseca, C.M., Fleming, P.J., Zitzler, E., Deb, K., Thiele, L., eds.: Conference on Evolutionary Multi-Criterion Optimization (EMO 2003). Volume 2632 of Lecture Notes in Computer Science., Berlin, Springer (2003) 494–508
24. Lehmann, E., D'Abrera, H.: Nonparametrics: Statistical Methods Based on Ranks. Prentice-Hall, Englewood Cliffs, NJ (1998)
25. Jensen, M.C.: The performance of mutual funds in the period 1945 - 1964. Journal of Finance **23** (May 1968) 383–417
26. Sharpe, W.: Capital assets prices: A theory of market equilibrium under conditions of risk. Journal of Finance **19** (1964) 425–442