

Exploring the (Efficient) Frontiers of Portfolio Optimization

Matthew J. Craven

Centre for Mathematical Sciences, Plymouth University
Plymouth PL4 8AA, UK
matthew.craven@plymouth.ac.uk

David I. Graham

Centre for Mathematical Sciences, Plymouth University
Plymouth PL4 8AA, UK
d.graham@plymouth.ac.uk

ABSTRACT

The cardinality-constrained portfolio optimization problem is NP-hard. Its Pareto front (or the Efficient Frontier - EF) is usually calculated by stochastic algorithms, including EAs. However, in certain cases the EF may be decomposed into a union of sub-EFs. In this work we propose a systematic process of excluding sub-EFs dominated by others, enabling us to calculate non-dominated sub-EFs. We then calculate whole EFs to a high degree of accuracy for small cardinalities, providing an alternative to EAs in those cases. We can use also this to provide insight into EAs on the problem.

CCS CONCEPTS

•Theory of computation → Optimization with randomized search heuristics; Quadratic programming; •Applied computing → Multi-criterion optimization and decision-making;

KEYWORDS

Portfolio optimization, Pareto front, quadratic programming, evolutionary algorithms, multi-objective problem

ACM Reference format:

Matthew J. Craven and David I. Graham. 2017. Exploring the (Efficient) Frontiers of Portfolio Optimization. In *Proceedings of GECCO '17 Companion*, Berlin, Germany, July 15-19, 2017, 2 pages.
DOI: <http://dx.doi.org/10.1145/3067695.3082036>

1 INTRODUCTION

The portfolio optimization problem [5] is an important one in finance. Given the returns of several assets over a period of time, an investor must decide in which proportions and in which assets to invest. The usual assumptions are that the investor wishes to maximize return and minimize risk within a given investment period. The Markowitz problem, i.e., no additional investment constraints are imposed other than all proportions being non-negative and summing to one, may be solved by quadratic programming (QP). To more accurately reflect real-world practice, additional constraints are commonly imposed. These conditions include cardinality constraints (restricting the number of assets in which we may invest from the set of all assets) and minimum proportions (if an asset is bought then a minimum proportion must be bought). However, addition of such constraints turns the problem into an NP-hard problem. In this work, we investigate the portfolio optimization problem under a cardinality constraint. That is, we invest in a “sub-portfolio”

of k of n assets, each asset having a proportion (weight) w_i of our capital. We write a portfolio as a vector of weights $\mathbf{w} = (w_1, w_2, \dots, w_n)$ with sum 1. Each \mathbf{w} has a mean return \bar{R} (from all returns in the investment using those weights) and a risk r (usually the standard deviation of returns). The standard expression of this problem is as a two-objective problem: we wish to minimize risk r subject to

$$\begin{aligned} \bar{R} &\geq R_0 \\ \sum_{i=1}^n w_i &= 1 \\ w_i &\geq 0 \text{ for } i = 1, \dots, n \\ \#\{j : w_j \neq 0\} &\leq k \end{aligned} \quad (1)$$

Equivalently, mean returns \bar{R} may be maximized subject to $r \leq r_0$ where r_0 is a given level of risk. Due to its NP-hardness, solving this problem is commonly done via a stochastic or evolutionary algorithm (EA) [3]. An EA is an optimizer which attempts to minimize or maximize one or more objectives related to a given problem. Many problems have more than one objective and, through use of parameterization or vectorization, may be converted to problems with fewer objectives (although some information may be lost). In this spirit, the portfolio problem under a cardinality constraint has the above expression (1) but may be expressed as a single-objective problem by substituting the objective with the following:

Minimize the function $\lambda r - (1 - \lambda)\bar{R}$ for a parameter λ subject to the weight constraints of (1).

With respect to problem instances, we use instances (D1)–(D5) of the benchmark OR-Library [1]. These are portfolio optimisation problems using, respectively, 31, 85, 89, 98 and 225 assets.

The Pareto front of (1) is referred to as the Cardinality Constrained Efficient Frontier (CCEF; see, e.g., Fig. 2). However, in the 1-objective case of the problem (parameterized by λ) the CCEF is not contiguous but instead contains “gaps” due to the parameterization [3]. In the 2-objective case there are no such gaps (but such CCEFs can be ragged). By definition, the CCEF contains all the non-dominated sub-portfolios (vectors \mathbf{w} with $n - k$ weights being zero such that if there are two portfolios \mathbf{w}, \mathbf{w}' then \mathbf{w} dominates \mathbf{w}' if the EF for \mathbf{w}' is both below and to the right of that of \mathbf{w}). In the 2-objective case, the full CCEF may be produced by considering optimal portfolios for R_0 in the range of minimal to maximal return. This, for one objective, is the finding of optimal portfolios for $\lambda \in [0, 1]$. Note that we do not compare runtimes with EAs in this work, as the purpose is to provide a method finding exact solutions which may be used as benchmark optimal solutions for an EA.

2 SIEVING AND SUB-EFS

We re-arrange returns and covariance data in order of asset returns (with highest returning asset corresponding to row / column 1). Generate the EF for the sub-portfolio consisting of (renumbered) assets 1, 2, \dots , k . Then for each of the $\binom{n}{k} - 1$ remaining sub-portfolios,

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '17 Companion, Berlin, Germany

© 2017 Copyright held by the owner/author(s). 978-1-4503-4939-0/17/07...\$15.00
DOI: <http://dx.doi.org/10.1145/3067695.3082036>

we test whether the EF to be generated by this sub-portfolio would be dominated. We do this by calculating three points on a “test EF” - one to give a maximum return \bar{R}_{\max} (corresponding to $\lambda = 0$), one to minimize the risk r_{\min} (at $\lambda = 1$) and another point with coordinate $(r_{\min}, \bar{R}_{\max})$. The test EF is dominated if, for all three points (r_i, \bar{R}_i) on the test EF, $r_j \geq r_i$ implies $\bar{R}_j > \bar{R}_i$ for every point (r_j, \bar{R}_j) on each previously computed non-dominated EF. If the test EF is not dominated in this way, calculate and store the full EF for this sub-portfolio (we use the Maple QP solver). The points on the non-dominated sub-EFs calculated for dataset (D1) of [1] are shown as blue circles in Fig. 1. At the end of this process we pool all data for the calculated sub-EFs. On the pooled data, we now perform a sieving process to determine whether individual sub-EF points are dominated. The final CCEF is the set of all non-dominated points (shown in orange in Fig. 1).

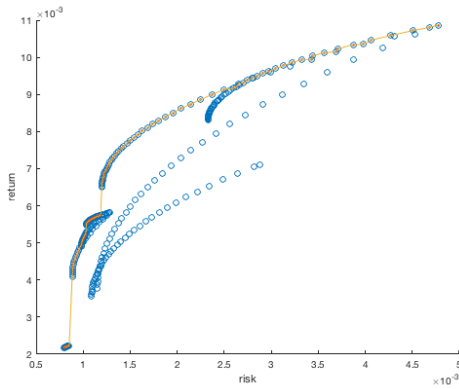


Figure 1: Sieved efficient frontier composed of non-dominated sub-portfolios for the (D1) dataset with $k = 2$.

Fig. 2 shows the CCEF generated by [2]. Comparing with Fig. 1, we see the final computed CCEF coincides perfectly with that generated by the authors except for the low risk/return region of the CCEF where our method slightly out-performs [2] by finding points whose risk is smaller than any of the points generated by [2]. [Note, the raw CCEF generated by the authors is ragged in the sense that it is non-monotonic and includes individual points which are dominated by other points. This is because the authors fix a return level and then solve a single-objective problem to find the minimum risk corresponding to that return. The CCEF shown here has used a sifting process to remove all of the dominated points and is monotonic; i.e., return increases with risk.]

Table 1 shows the number of non-CCEF-dominated EF’s found for the OR-library datasets with sub-portfolios of various sizes. In all cases, it is clear that the vast majority of sub-portfolios are CCEF-dominated and full sub-EFs must be computed only for a very small number of portfolios. This means that cases where the number of possible sub-portfolios is vast remain computable within a reasonable time. For example, there are almost 8 million sub-portfolios for the case (D1), $k=8$. Only four of these are non-CCEF-dominated and computation was completed within a day.

3 RESULTS AND CONCLUSIONS

In [4] we attempted to solve the 1-objective problem for the datasets in [1] with an EA. As the problem is especially difficult for small

values of k (e.g., $k = 2$ is the most difficult) the EA runtime was large. However, moderate values $k \geq 10$ were solved efficiently by the EA. The present work gives an alternative method for solving the problem with a combined sieve-QP methodology. The number of combinations of k assets from n (i.e. $\binom{n}{k}$) is small for small k (Table 1, third column) and large for the same moderate k above. This illustrates a pleasing duality: the sieve-QP methodology may be used for small k values while an EA may be used for other k values. The method is exact providing the QP procedure used converges, and we have evidence that it is also effective for distinct risk measures (in which case a non-QP solver may be necessary).

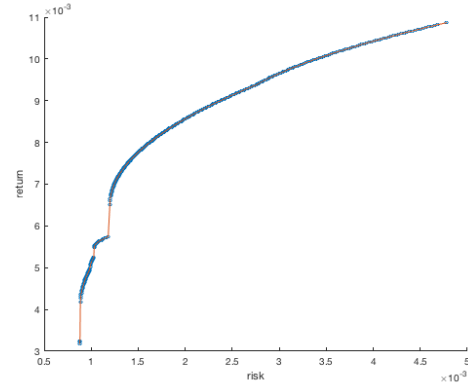


Figure 2: Sieved CCEF of [2] for the instance of Fig. 1.

The analysis of generated CCEFs and dominated portfolios drives insight into EA processes. Our computations suggest that the sub-EFs are local minima of the cost function of common EAs on the problem, i.e., an EA picks out the sub-optimal sub-EFs discarded by the above procedure through cost minimisation (attempting to move to sub-EFs, i.e., solutions, closer to the CCEF). We shall investigate this further.

Table 1: The number of sub-portfolios and non-dominated sub-EFs for various k for each dataset in [1].

Dataset	k	# Sub-portfolios	# Non-Dom. Sub-EFs
D1	2	465	8
	4	31465	7
	6	736281	6
	8	7888725	4
D2	2	3570	7
	4	2024785	6
D3	2	3916	5
	4	2441616	6
D4	2	4753	10
	4	3612280	9
D5	2	25200	7
	3	1873200	17

REFERENCES

- [1] J. E. Beasley. 2000. OR-Library. people.brunel.ac.uk/~mastjb/jeb/orlib/portinfo.html (2000).
- [2] F. Cesarone, A. Scozzari, and F. Tardella. 2010. Solutions to Portfolio Problems in Beasley’s OR-Library. (2010). host.uniroma3.it/docenti/cesarone/datasetsw3_tardella.html.
- [3] T. J. Chang, N. Meade, J. E. Beasley, and Y. M. Sharaiha. 2000. Heuristics for Cardinality Constrained Portfolio Optimisation. *Comp. & Oper. Res.* 27, 13 (2000), 1271–1302.
- [4] Matthew J Craven and David I Graham. 2017. Parallel Stochastic Solution of Portfolio Optimisation Problems with Cardinality Constraints. *submitted* (2017).
- [5] H. Markowitz. 1952. Portfolio Selection. *J. Finance* 7, 1 (1952), 77–91.