

Large-Scale Portfolio Optimisation with Heuristics¹

Manfred Gilli²

Department of Econometrics, University of Geneva, e-mail: Manfred.Gilli@unige.ch

Enrico Schumann

Department of Econometrics, University of Geneva, e-mail: Enrico.Schumann@unige.ch

Abstract: Heuristic optimisation techniques allow to optimise financial portfolios with respect to different objective functions and constraints, essentially without any restrictions on their functional form. Still, these methods are not widely applied in practice. One reason for this slow acceptance is the fact that heuristics do not provide the ‘optimal’ solution, but only a stochastic approximation of the optimum. For a given problem, the quality of this approximation depends on the chosen method, but also on the amount of computational resources spent (eg, the number of iterations): more iterations lead (on average) to a better solution. In this paper, we investigate this convergence behaviour for three different heuristics: Differential Evolution, Particle Swarm Optimisation, and Threshold Accepting. Particular emphasis is put on the dependence of the solutions’ quality on the problem size, thus we test these heuristics in large-scale settings with hundreds or thousands of assets, and thousands of scenarios.

Keywords: Portfolio Optimisation, Heuristic Optimisation, Differential Evolution, Particle Swarm Optimisation, Threshold Accepting

1. Introduction

The popularity of mean–variance optimisation (Markowitz, 1952) in finance stems to a considerable extent from the ease with which such models can be solved numerically. The model requires that investors decide on their portfolios solely on the basis of the first two moments of return distribution. Given the non-Gaussian nature of financial time series (Cont, 2001), alternative selection criteria have been proposed that take into account empirical regularities like asymmetric return distributions and ‘fat tails’. But unfortunately, these criteria are rarely used in realistic portfolio selection problems, mainly because the optimisation becomes much more difficult. Heuristics have been shown to work well for such problems that are completely infeasible for standard techniques like gradient-based methods (for heuristics in general, see Michalewicz and Fogel (2004); for financial applications, see Maringer (2005)).

In this paper, we will discuss the application of heuristics to such portfolio selection problems, in particular large-scale problems. While mathematically the properties of a given optimisation problem may not change with its size, the same does not hold true numeri-

¹The authors gratefully acknowledge financial support from the EU Commission through MRTN-CT-2006-034270 COMISEF.

²Address of correspondence: Corresponding author: Department of Econometrics, University of Geneva, Bd du Pont d’Arve 40, 1211 Geneva 4, Switzerland. Tel.: +41 22 379 8222; fax: +41 22 379 8299.

cally. The paper analyses the performance of three particular methods, Differential Evolution, Particle Swarm Optimisation, and Thresholds Accepting, and how their performance changes with the size of the problem.

2. Methodology and techniques

2.1 The model

We assume that there are n_A risky assets available, with current prices collected in a vector p_0 . We are endowed with an initial wealth v_0 , and wish to select a portfolio $x = [x_1 \ x_2 \ \dots \ x_{n_A}]'$ of the given assets. The chosen portfolio is held for one period, from now (time 0) to time T . End-of-period wealth is given by

$$v_T = x' p_T,$$

where the vector p_T holds the asset prices at T . Since these prices are not known at the time when the portfolio is formed, v_T will be a random variable, following some unknown distribution (that depends on x).

Then our optimisation problem can be written as

$$\begin{aligned} \min_x \quad & \Phi, \\ & x_j^{\inf} \leq x_j \leq x_j^{\sup} \quad j \in \mathcal{J}, \\ & K_{\inf} \leq \#\{\mathcal{J}\} \leq K_{\sup}. \end{aligned}$$

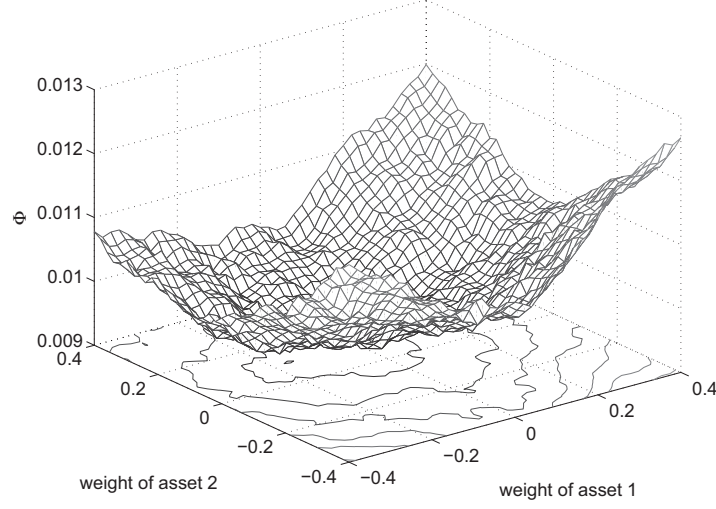
Φ is the particular objective function chosen, for instance a downside risk measure, or a risk–reward ratio (Gilli and Schumann, 2009). $\mathcal{J}$ is the set of assets in the portfolio, $x_j^{\inf}$ and $x_j^{\sup}$ are minimum and maximum holding sizes, respectively, for the assets included in $\mathcal{J}$. $K_{\inf}$ and $K_{\sup}$ are cardinality constraints that set a minimum and maximum number of assets in $\mathcal{J}$. We also include other kinds of constraints like exposure limits to specific risk factors, or sector constraints.

2.2 Techniques

The model outlined above combines combinatorial (pick between $K_{\inf}$ and $K_{\sup}$ assets from a set of n_A assets) and continuous elements (the objective function, maximum holding size constraints); it can generally not be solved with standard methods since it is not convex and exhibits many local minima. As an example, Figure 1 shows the search space (ie, the mapping from portfolio weights into the objective functions) for a three asset problem with the objective to minimise the Value-at-Risk.

Heuristics are explicitly designed to escape local minima and thus are appropriate methods for such problems. We compare three different techniques: Differential Evolution (DE; Storn and Price (1997)), Particle Swarm Optimisation (PS; Eberhart and Kennedy (1995)), and Threshold Accepting (TA; Dueck and Scheuer (1990)). The main emphasis of the paper is on the performance of these methods for large-scale instances of the problem.

Figure 1: *The search space for a Value-at-Risk optimisation (3 assets).*



2.3 Data

Since we focus on the efficacy of the optimisation techniques (not on the empirical, out-of-sample performance of the selected portfolios), we can create our data through simulation. This has the advantage that we can vary the problem size in terms of observations and numbers of assets, and compare the respective performance of the different techniques. We characterise every asset's return by a linear factor model,

$$r^i = \beta_0^i + \beta_1^i f_1 + \dots + \beta_{n_F}^i f_{n_F} + \epsilon^i = \beta^i f + \epsilon^i.$$

'Factor' model is to be understood not in the sense of asset-pricing models, but from the perspective of the practitioner who models the assets' returns, in particular their cross-sectional dependencies. Thus, the i th asset is characterised by its factor loadings, β^i , and by the residuals ϵ^i that capture the remaining variation. To simulate returns, we draw factor and residual returns from specified distributions.

We also add binary (zero-one) variables for every asset. These may be used to model observable characteristics, for instance whether a stock belongs to a certain industry or comes from a specific country. These binary variables will not be used in the return-generation procedure, but they allow the implementation of practical constraints. We may, for example, include several cardinality restrictions (no more than n stocks from country z), or use them to aggregate exposures to specific industry sectors.

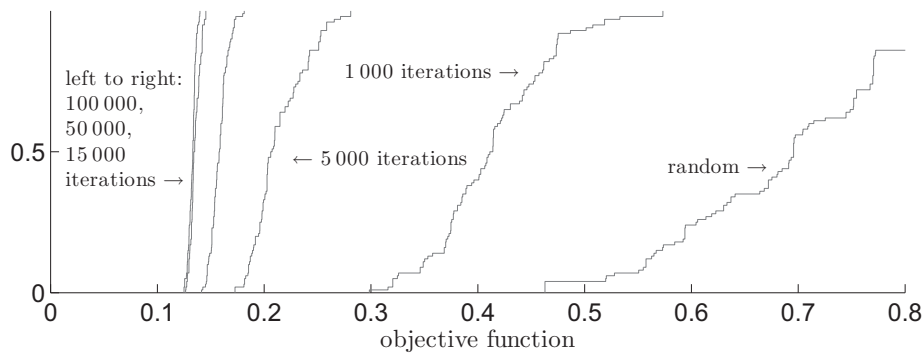
3. Convergence results

Heuristic methods are, with only few exceptions, stochastic algorithms. Every restart of the algorithm can thus be regarded as the realisation of a random variable with some unknown distribution. For a given problem, this distribution will depend on the method chosen: some techniques may be more appropriate than others, and thus give less variable and better results. The distribution will also be influenced by the amount of computational

resources (ie, the number of iterations) spent. Since heuristics can move away from local minima, allowing more iterations generally leads to better solutions on average (in fact, with iterations going to infinity, the distribution collapses to a single point, the global minimum).

We investigate this convergence for our three methods. Figure 2 shows an example for TA for a relatively small problem with 300 assets and 5 000 scenarios (observations). The objective here was to minimise the downside semi-variance. Every distribution is computed from 100 restarts of the algorithm. As can be seen, with increasing numbers of iterations, the distribution move to the left and becomes steeper (less variable).

Figure 2: *Convergence of TA (n_A : 300, n_O : 5 000).*



We compare this convergence behaviour for all our methods, in particular when the problem size increases and when more constraints are added.

References

- Cont R. (2001) Empirical properties of asset returns: Stylized facts and statistical issues, *Quantitative Finance*, 1, 223–236.
- Dueck G. and Scheuer T. (1990) Threshold Accepting. A General Purpose Optimization Algorithm Superior to Simulated Annealing, *Journal of Computational Physics*, 90, 1, 161–175.
- Eberhart R.C. and Kennedy J. (1995) A new optimizer using particle swarm theory, in: *Proceedings of the Sixth International Symposium on Micromachine and Human Science*, Nagoya, Japan, 39–43.
- Gilli M. and Schumann E. (2009) An Empirical Analysis of Alternative Portfolio Selection Criteria, *Swiss Finance Institute Research Paper No. 09-06*.
- Maringer D. (2005) *Portfolio Management with Heuristic Optimization*, Springer.
- Markowitz H.M. (1952) Portfolio selection, *Journal of Finance*, 7, 1, 77–91.
- Michalewicz Z. and Fogel D.B. (2004) *How to Solve it: Modern Heuristics*, Springer.
- Storn R. and Price K. (1997) Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, 11, 341–359.