Portfolio selection with an alternative measure of risk: computational performances of particle swarm optimization and genetic algorithms

Marco Corazza, Giovanni Fasano and Riccardo Gusso

Abstract In the classical model for portfolio selection the risk is measured by the variance of returns. Recently several alternative measures of risk have been proposed. In this contribution we focus on a class of measures that uses information contained both in lower and in upper tail of the distribution of the returns. We consider a nonlinear mixed-integer portfolio selection model which takes into account several constraints used in fund management practice. The latter problem is NP-hard in general, and exact algorithms for its minimization, which are both effective and efficient, are still sought at present. Thus, to approximately solve this model we experience the heuristics Particle Swarm Optimization (PSO) and we compare the performances of this methodology with respect to another well-known heuristic technique for optimization problems, that is Genetic Algorithms (GA).

Keywords: Portfolio selection problem, measures of risk, constrained optimization, evolutionary optimization, Particle Swarm Optimization, Genetic Algorithms.

1 Introduction to PSO

Particle Swarm Optimization is an iterative heuristics for the solution of nonlinear global optimization problems [10]. It is based on a biological paradigm, which is inspired by the flight of birds in a flock, looking for food. Every member of the

Marco Corazza
Department of Economics and Advanced School of Economics – Ca' Foscari University of Venice – Fondamenta San Giobbe, Sestiere Cannaregio 873, 30121 Venezia (Italy), e-mail: corazza@unive.it.

Giovanni Fasano
Department of Management – Ca' Foscari University of Venice – Fondamenta San Giobbe, Sestiere Cannaregio 873, 30121 Venezia (Italy), e-mail: fasano@unive.it.

Riccardo Gusso
Department of Economics – Ca' Foscari University of Venice – Fondamenta San Giobbe, Sestiere Cannaregio 873, 30121 Venezia (Italy), e-mail: rgusso@unive.it.

flock explores the search area keeping memory of its best position reached so far, and it exchanges this information with its neighbors.

In its mathematical counterpart the paradigm of a flying flock may be formulated as follows: find a global minimum (best global position) in a nonlinear and nonconvex minimization problem. Every member of the swarm (namely a *particle*) represents a possible solution of the minimization problem, and it is initially positioned randomly in the feasible set of the problem. Every particle is also initially assigned with a random *velocity* which is used to determine its initial direction of movement.

The overall PSO algorithm with M particles , as in the version with inertia weight proposed in [13], works as follows in a minimization problem:

1. Set $k = 1$ and evaluate $f(\mathbf{x}_j^k)$ for $j = 1,\dots,M$. Set $pbest_j = +\infty$ for $j = 1,\dots,M$.
2. If $f(\mathbf{x}_j^k) < pbest_j$ then set $\mathbf{p}_j = \mathbf{x}_j^k$ and $pbest_j = f(\mathbf{x}_j^k)$.
3. Update position and velocity of the j-th particle, $j = 1,\dots,M$, as

$$\mathbf{v}_j^{k+1} = w^{k+1}\mathbf{v}_j^k + \mathbf{U}_{\phi_1} \otimes (\mathbf{p}_j - \mathbf{x}_j^k) + \mathbf{U}_{\phi_2} \otimes (\mathbf{p}_{g(j)} - \mathbf{x}_j^k) \tag{1}$$

$$\mathbf{x}_j^{k+1} = \mathbf{x}_j^k + \mathbf{v}_j^{k+1} \tag{2}$$

   where $\mathbf{U}_{\phi_1}, \mathbf{U}_{\phi_2} \in {}^d$ and their components are uniformly randomly distributed in $[0,\phi_1]$ and $[0,\phi_2]$ respectively.
4. If a *convergence test* is not satisfied then go to 1.

The symbol $\otimes$ denotes component-wise product and $\mathbf{p}_{g(j)}$ is the best position in a neighborhood of the j-th particle. The specification of the neighborhood topology is then a choice to set. In our implementation we have considered the so called *gbest* topology, that is $g(j) = g$ for every $j = 1,\dots,M$, and g is the index of the best particle in the whole swarm. The value of the inertia weight w^k, a parameter that forces the convergence of the swarm to single solution and prevents the "explosion" of the particles' trajectories in the search space, is generally linearly decreasing with the number of steps, i.e.

$$w^k = w_{max} + \frac{w_{min} - w_{max}}{K} k \, . \tag{3}$$

In this work we have used the most common values for w_{max} and w_{min} found in the literature, that are respectively 0.9 and 0.4, while K is the maximum number of steps allowed.

2 Portfolio Selection and Risk Measures

The basic idea in the portfolio selection problem is to select stocks in order to maximize the portfolio performance and at the same time to minimize its risk. This implies that for a formal approach to the latter problem, a correct definition of *performance* and *risk* of the portfolio is required. While there is a general agreement about the measurement of performance by the expected value of the future return of the portfolio, the discussion regarding an adequate measure of risk is still open.

In the classical approach, since the work of Markowitz [11], variance is used to measure risk, but this has one major shortcoming: it leads to optimal investment decisions only if investment returns are elliptically distributed or if the utility function of investors is quadratic. This consideration has opened the way for the research on alternative measures of risk, and recently there has been a growing interest for the so called *coherent risk measures* introduced in [1].

In [4] Chen and Wang have investigated the possibility of building a new class of coherent risk measures, by combining upper and lower moments of different orders. This approach seems to have several advantages with respect to others considered so far. Indeed, on one hand these measures better couple with non normal distributions than ones based only on first order moments. On the other hand, they better reflect investors' risk attitude, for at least a couple of reasons. First they are less affected by estimation risk than measures that use only information from the lower part of the return distribution. Moreover, according with the conclusions presented in [4], their use in the portfolio selection problem allows for more realistic and robust results, compared with the ones obtained using CVaR.

In this contribution we use this class of risk measures for a portfolio selection problem similar to the one considered in [4], with the addition of the cardinality constraints, which yield a final model in the class of nonlinear mixed-integer programming problems. For the latter scheme (which is an NP-hard problem [12]) at present there are not both efficient and effective algorithms as for the problem considered in [4]: this motivates the possible introduction of evolutionary heuristic methodologies as PSO.

2.1 The portfolio selection model

Let X be a real valued random variable defined on a probability space $(\Omega, \mathscr{F},)$, and let us denote $\|X\|_p = ([|X|^p])^{1/p}$, $p \in [1, +\infty[$, where $[\cdot]$ indicates the expected value of a random variable. Then, the measures of risk introduced in [4] are defined as:

$$\rho_{a,p}(X) = a\|(X - [X])^+\|_1 + (1-a)\|(X - [X])^-\|_p - [X], \tag{4}$$

where $a \in [0,1]$, $X^- = \max\{-X, 0\}$ and $X^+ = (-X)^-$.

For a and p fixed, any risk measure of this class is a coherent risk measure (see [7]): for a proof of this and a detailed description of its properties we refer the reader to [4]. We only remark here that $\rho_{a,p}$ is non-decreasing with respect to p and non-increasing with respect to a. Thus, the value of these parameters can be adjusted to reflect different attitudes of the investors towards risk.

The portfolio selection model we consider is the following one: suppose we have N assets to choose from, and for $i = 1, \ldots, N$ let $x_i \in$ be the weight of asset i in the portfolio, with $X^T = (x_1 \cdots x_N)$. Let $z_i \in \{0, 1\}$ with $Z^T = (z_1 \cdots z_N)$ be a binary variable, such that $z_i = 1$ if the asset i is included in the portfolio, $z_i = 0$ otherwise. Moreover, for $i = 1, \dots, N$, let r_i be a real valued random variable that represents the return of asset i, with $\hat{r}_i$ its expected value, i.e. $\hat{r}_i = [r_i]$. Then, the random variable $R \in$ that represents the return of the whole portfolio can be expressed as $R = \sum_{i=1}^{N} x_i r_i$, with expected value $\hat{R} = \sum_{i=1}^{N} x_i \hat{r}_i$.

Then, our overall portfolio selection problem can be written as follows:

$$\begin{aligned} \min_{X,Z}\ & \rho_{a,p}(R) \\ \text{s.t. } & \hat{R} \geq l \\ & \sum_{i=1}^{N} x_i = 1 \\ & K_d \leq \sum_{i=1}^{N} z_i \leq K_u \\ & z_i d \leq x_i \leq z_i u, \qquad i = 1, \dots, N, \\ & z_i(z_i - 1) = 0, \qquad i = 1, \dots, N. \end{aligned} \tag{5}$$

The first constraint in (5) represents the the minimum desirable expected return l of the portfolio, while the second one is the usual budget constraint. Then we have the cardinality constraint: we neither select a too small subset of our assets (K_d) nor a too large one (K_u). The latter choice summarizes a quite common problem for a fund manager, who has to build a portfolio by choosing from several hundreds of assets. Moreover, we require that any of the selected assets x_i must not constitute a too large or too small fraction of the portfolio (i.e. $z_i d \leq x_i \leq z_i u$, where d and u are positive parameters, with $d \leq u$). The last N constraints are introduced to model the relations $z_i \in \{0, 1\}$, $i = 1, \dots, N$.

Of course, (5) is a reformulation of a *nonlinear and nonconvex mixed-integer problem*, where the constraints $z_i \in \{0, 1\}$, $i = 1, \dots, N$, are replaced by the *relaxations* $z_i(1 - z_i) = 0$, $i = 1, \dots, N$. Detecting precise solutions of (5) may be heavily time consuming in case exact methods are adopted.

3 Optimization using PSO and GA: reformulation of the portfolio selection problem

Originally PSO was conceived for unconstrained problems. Thus, in general using PSO formulae (1)-(2), when constraints are included in the formulation, is improper. Indeed, in the latter case the PSO algorithm cannot prevent from generating infeasible particles' positions, unless specific adjustments are adopted. When constraints are included, different strategies were proposed in the literature (see also [2]) to ensure that at any step of PSO, feasible positions are generated. Most of them involve repositioning of the particles, as for example the *bumping* and *random positioning* strategies proposed in [15]. In this paper we decided to use PSO as in its original formulation, so we have reformulated our problem into an unconstrained one, using the nondifferentiable ℓ_1 *penalty function* method described in [14, 8]. Our reformulation of (5) (which has $N+1$ equality constraints and $2N+3$ inequality constraints) is given by

$$\min_{X,Z} P(X,Z;\varepsilon)$$

and uses the nondifferentiable penalty function

$$\begin{aligned} P(X,Z;\varepsilon) \;=\; & \rho_{a,p}(R) + \frac{1}{\varepsilon}\Bigg[\max\{0, l-\hat{R}\} + \left|\textstyle\sum_{i=1}^{N} x_i - 1\right| + \\ & + \max\left\{0, K_d - \textstyle\sum_{i=1}^{N} z_i\right\} + \max\left\{0, \textstyle\sum_{i=1}^{N} z_i - K_u\right\} \\ & + \textstyle\sum_{i=1}^{N} \max\{0, z_i d - x_i\} \\ & + \textstyle\sum_{i=1}^{N} \max\{0, x_i - z_i u\} + \textstyle\sum_{i=1}^{N} |z_i(1-z_i)|\Bigg] \end{aligned} \tag{6}$$

and ε is the penalty parameter. The correct choice of ε ensures the correspondence between the solutions of problems (6) and (5) (see also [6]). Of course, since PSO is a heuristics, the minimization of the penalty function $P(X,Z;\varepsilon)$ theoretically does not ensure that a global minimum of (5) is detected. Nevertheless, PSO often provides a suitable compromise between the performance (i.e. a satisfactory estimate of a global minimizer for (5)) and the computational cost. To analyze the performance of PSO we compare it with another well known evolutionary heuristic methodology for optimization problems, that is a *genetic algorithm* (GA) in its standard form, that is starting from an initial population of solutions we generate a new one using the following three steps: tournament, basic crossover and basic mutation. For sake of brevity we refer the reader to [9] for more details on GA.

4 Numerical results

In this section we briefly report the conclusions of the numerical results we have obtained (for further results and references see also [5]).

As input data we have used the time series of the daily close prices of the 32 assets belonging to Italian FTSE MIB index from January 2003 to May 2009. Using the same idea of [4] we have estimated the risk measure for any portfolio X as

$$\begin{aligned} \rho_{a,p}(R) = \frac{a}{T}\left[\sum_{t=1}^{T} \left(\sum_{i=1}^{N}(r_{i,t}-\hat{r}_i)x_i\right)^{+}\right] \\ + (1-a)\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\left(\sum_{i=1}^{N}(r_{i,t}-\hat{r}_i)x_i\right)^{-}\right]^{p}\right\}^{\frac{1}{p}}, \end{aligned}$$

where $\hat{r}_i$ is estimated using the historical data, that is

$$\hat{r}_i = \frac{1}{T}\sum_{t=1}^{T} r_{i,t}.$$

To reflect a realistic problem of portfolio selection, we set the values $d = 0.05$ and $u = 0.20$ in (6). For the cardinality constraint we have set $K_d = 5$, while we have considered two different values for K_u: $K_u = 20$ and $K_u = 10$. The PSO and GA algorithms to solve problem (6) has been implemented in MATLAB 7, and the experiments have been performed on a workstation Acer Aspire M1610 with an Intel Core 2 Duo E4500 processor.

We stopped PSO iterations when either of the following stopping criteria was satisfied:

a) the maximum number of 10000 steps was outreached;
b) $|fbest^{k+1} - fbest^{k}| < 10^{-8}$ for 2000 consecutive steps, where $fbest = f(\mathbf{p}_g)$ is the current best value of the fitness function $f = P(X,Z;\varepsilon)$.

After some preliminary tests, aiming to use values for the parameters as standard as possible in the literature, we selected the values $\varepsilon = 10^{-6}$ and $M = 50$.

We solved the portfolio selection problems for different values of the parameters of the risk measure $\rho_{a,p}$, and K_u, considering one year data of daily returns of different time periods. For every combinations of the parameters and the data-set, we did first 50 runs of the algorithm, each with different random initial positions and velocities. We then iterated the procedure in the following way: we did other 50 runs of the algorithm, with again random initial velocities for all particles, but we used the 50 global best positions found in the previous phase as initial positions. At the end of this second phase we obtained convergence to the same global best position for each run (in general not corresponding to the best position of the previous 50 ones) and we assumed this to be the global minimum (X^*,Z^*) of the optimization problem.

We remark that the monotonicity properties expected by theoretical results ([4, Theorem 2.3]) were respected by the results found using PSO, also with $K_u = 10$. This is shown in Table 1 and Table 2. We also observe that the diversification of the portfolio, measured by the number of assets, is decreasing with a and increasing with p, and this is consistent with the different attitudes towards risk expressed by the values of these parameters. The same considerations apply using data from different time periods.

Table 1 Monotonicity of $\rho_{a,p}(X^*)$ for $p = 2$ and different values of a and K_u, with one year data from 2003-04.

	$a=0$	$a=0.25$	$a=0.5$	$a=0.75$	$a=1$
$\rho_{a,2;K_u=20}$	0.004962	0.004667	0.003560	0.002816	0.002165
N. of assets	20	18	17	16	15
$\rho_{a,2;K_u=10}$	0.004968	0.004748	0.003619	0.002934	0.002372
N. of assets	9	10	9	9	9

Table 2 Monotonicity of $\rho_{a,p}(X^*)$ for $a = 0.5$ and different values of p and K_u, with one year data from 2003-04.

	$p=1$	$p=2$	$p=5$
$\rho_{0.5,p;K_u=20}$	0.002024	0.00356	0.006787
N. of assets	15	18	19
$\rho_{0.5,p;K_u=10}$	0.00209	0.003619	0.00697
N. of assets	8	10	10

To analyze the performance of PSO with respect to GA, since the results in terms of the risk measure are approximately the same, we compared the standard deviations of the optimal risk measure in the first 50 runs of the two algorithms, in order to investigate the consistency of the algorithms, that is the capability of the two methods to converge to the same solution in each run. We observed a little better performance of GA in this respect, especially in the case $K_u = 10$, but this has a cost: the average computational time in seconds is then approximately 10 times larger. An example of the results obtained is reported in Table 3.

Table 3 Average standard deviation and computational time of 50 runs of PSO and GA for $p = 2$, $K_u = 10$ and different values of p.

	$a=0$	$a=0.25$	$a=0.5$	$a=0.75$	$a=1$
$\sigma(PSO)$	0.0904%	0.1026%	0.0727%	0.0657%	0.0315%
$\sigma(GA)$	0.0751%	0.0701%	0.0636%	0.0529%	0.0286%
$\bar{t}(PSO)$	30.04	31.25	29.87	30.35	32.03
$\bar{t}(GA)$	315.27	298.82	308.12	330.23	321.73

In order to analyze the financial meaning of the portfolios obtained, we used PSO to solve another portfolio selection problem, using variance as measure of risk, and keeping the same set of constraints of problem (5). By comparing the diversification of the two portfolios obtained, it appears that when the cardinality constraint is in its weaker form, that is $K_u = 20$, the diversification obtained using $\rho_{a,p}$ is higher than using variance, and it is increasing with p. This is also consistent with the results obtained in [4], where the cardinality constraint was not explicitly introduced, and the comparison was made with respect to CVaR.

5 Conclusions

The results obtained suggest that when challenging nonlinear and nonconvex mixed-integer reformulations of portfolio selection are considered, including complex objective function landscapes and a set of constraints, then PSO provides a satisfactory compromise between the performance and the computational workload required. The latter conclusion comes up from our experience, by comparing PSO with GA, when the dimensionality of the problem is high. More investigation is needed to check the dependence of the performance of PSO with respect to the initial position and velocities of the particles (see [3]) and to a different strategy for handling the constraints of the problem.

References

1. Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Mathematical Finance **9**(3), 203–228 (1999)
2. Blackwell, T., Kennedy, J., Poli, R.: Particle swarm optimization - an overview. Swarm Intelligence **1**(1) (2007)
3. Campana, E.F., Fasano, G., Pinto, A.: Dynamic analysis for the selection of parameters and initial population, in particle swarm optimization. Journal of Global Optimization **48**(3), 347–397 (2010)
4. Chen, Z., Wang, Y.: Two-sided coherent risk measures and their application in realistic portfolio optimization. Journal of Banking and Finance **32**, 2667–2673 (2008)
5. Cura, T.: Particle swarm optimization approach to portfolio optimization. Nonlinear Analysis: Real World Applications **10**, 2396–2406 (2009)
6. Di Pillo, G., Grippo, L.: Exact penalty functions in constrained optimization. SIAM Journal on Control and Optimization **27**(6), 1333–1360 (1989). DOI 10.1137/0327068. URL http://link.aip.org/link/?SJC/27/1333/1
7. Fischer, T.: Risk capital allocation by coherent risk measures based on one-sided moments. Insurance: Mathematics and Economics **32**(1), 135–146 (2003)
8. Fletcher, R.: Practical methods of optimization. John Wiley & Sons, Glichester pp. 1942–1948 (1991)
9. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning, 1 edn. Addison-Wesley Professional (1989). URL http://www.worldcat.org/isbn/0201157675
10. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. Proceedings of the IEEE international conference on neural networks IV pp. 1942–1948 (1995)
11. Markowitz, H.M.: Portfolio selection. Journal of Finance **7**, 77–91 (1952)
12. Papadimitriou, C.H., Steiglitz, K.: Combinatorial optimization: Algorithms and complexity. Prentice-Hall, Inc., New Jersey pp. 1942–1948 (1982)
13. Shi, Y., Eberhart, R.: A modified particle swarm optimizer. In: Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, pp. 69–73 (1998). DOI 10.1109/ICEC.1998.699146. URL http://dx.doi.org/10.1109/ICEC.1998.699146
14. Zangwill, W.I.: Non-linear programming via penalty functions. Management Science **13**(5), 344–358 (1967). URL http://www.jstor.org/stable/2627851
15. Zhang, W., Xie, X.F., Bi, D.C.: Handling boundary constraints for numerical optimization by particle swarm flying in periodic search space. ArXiv Computer Science e-prints (2005)