Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

CrossMark

Particle Swarm Optimization with non-smooth penalty reformulation, for a complex portfolio selection problem

Marco Corazza [a,b,*], Giovanni Fasano [c,d], Riccardo Gusso [a]

[a] Department of Economics, Ca' Foscari University of Venice, Sestiere di Cannaregio 873, 30121 Venezia, Italy
[b] Advanced School of Economics in Venice, Sestiere di Cannaregio 873, 30121 Venezia, Italy
[c] Department of Management, Ca' Foscari University of Venice, Sestiere di Cannaregio 873, 30121 Venezia, Italy
[d] National Research Council-Maritime Research Centre (CNR-INSEAN), Via di Vallerano 139, 00128 Roma, Italy

ARTICLE INFO

ABSTRACT

In the classical model for portfolio selection the risk is measured by the variance of returns. It is well known that, if returns are not elliptically distributed, this may cause inaccurate investment decisions. To address this issue, 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 non-linear 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). Since PSO was originally conceived for unconstrained global optimization problems, we apply it to a novel reformulation of our mixed-integer model, where a standard exact penalty function is introduced.

© 2013 Published by Elsevier Inc.

1. Introduction

Making effective portfolio selection in real stock markets is a not-so-easy task for at least the following three reasons.

First, we need to gauge the risk by measures that both satisfy appropriate formal properties (namely coherence) and better couple with the non normal return distributions (which characterize the stock markets). Moreover, it should be desirable that these risk measures were suitably parameterized with respect to the investors' risk attitude. In other terms, we need personalized coherent risk measures that well fit a non Gaussian (financial) world.

Second, we have to take into account several practices and rules of the portfolio management industry that can affect the portfolio selection process (for instance, the use of bounds for the minimum and the maximum number of stocks to trade). Generally, such aspects are formalized in terms of constraints that very often yield NP-hard mathematical programming formulations.

Third, portfolio selection problems arising from the joint use of the considered risk measures, practices and rules are possibly highly nonlinear, nondifferentiable, nonconvex and mixed-integer, which contributes to yield NP-hardness. Therefore, the development of *ad hoc* solution approaches is usually needed to find the optimal solution of such problems. However, the portfolio management industry usually does not possess the mathematical knowledge and/or the research capabilities to handle such approaches. Furthermore, it could be not convenient for it to build up a team of external experts. As a

* Corresponding author at: Department of Economics, Ca' Foscari University of Venice, Sestiere di Cannaregio 873, 30121 Venezia, Italy.
E-mail addresses: corazza@unive.it (M. Corazza), fasano@unive.it (G. Fasano), rgusso@unive.it (R. Gusso).

consequence, part of the investors' demand might remain unsatisfied or (worse) satisfied by the use of inappropriate solution technologies. Therefore, we need a general tool encompassing a large variety of real portfolio selection problems.

In this paper we deal with such issues, and we propose a tool for managing each of them. Then, we perform a numerical experience on the resulting scheme, by applying it to the selection of large and complex portfolios. In particular, we tackle the above mentioned issues as follows.

First, as measure of risk of the portfolio returns we consider a recently proposed coherent risk measure, based on the combination of upper and lower moments of different orders of the returns distribution (see [7]). The latter measure shows to be able to effectively manage non Gaussian distributions of asset returns and to appropriately reflect different investors' risk attitudes (see Section 2 for details). In particular, given personal investors' weighting, it permits to take into account both the risk contained in the "bad tail" (the left one of the portfolio returns), and the chances contained in the "good tail" (the right one of the same portfolio). Further, to the best of our knowledge, apart from [7] this is one of the first applications of such risk measure to large and complex portfolios. In addition, we highlight that the portfolio selection problem considered in [7] is large but-not-complex, since the associated optimization problem is linearly constrained (i.e. relatively "easy" to solve). Conversely, our portfolio selection problem is large and complex, due to the presence of nonlinear mixed-integer constraints (see Section 2 for details) which make the associated optimization problem NP-hard (i.e. difficult to solve).

Second, as the professional practices and rules are concerned, following the indications of a north-eastern Italian company skilled in automatic financial trading systems, we adapt our analysis to the use of bounds for the minimum and the maximum number of stocks to trade.[1] Moreover, our model also includes the minimum and the maximum capital percentage to invest in each asset (see also [20] where the quantity to invest in each asset is a positive integer). These bounds are often of interest for the fund manager industry, in order to control (in an indirect way) the transaction costs. All these practices/rules are formalized in terms of constraints, that surely make the corresponding mathematical programming problem NP-hard (see Section 2 for details). Notice that the portfolio selection problem so arising is new in the specialized literature.

Third, our resulting selection problem, which takes into account also standard constraints (namely, the minimum return and the budget constraints) is nonlinear, nondifferentiable and mixed-integer. At present, for such mathematical programming problem (which is NP-hard, see [23]) both efficient and effective solution algorithms are still sought. Thus, in order to both investigate the numerical complexity of our portfolio selection problem and to provide a cheap and reliable solution, we adopt an exact penalty method (see [31,24,9,16]) combined with the Particle Swarm Optimization (PSO), a recently proposed bio-inspired population-based metaheuristic (see [17] as, likely, first contribution on it). In short, our solution approach runs as follows (see Section 3 for details):

- a standard exact penalty scheme transforms the considered mixed-integer portfolio selection problem into an equivalent nondifferentiable unconstrained minimization problem;
- then, as also the latter model is nonlinear, nondifferentiable and non convex, for its minimization a derivative-free algorithm is a possible solution method. Among the various approaches proposed in the specialized literature, we consider the PSO in order to approximately computing a global minimizer of the overall exact penalty-based model.

More generally, the choice of bio-inspired metaheuristics as global optimizers is also motivated by the fact that ≪[t]*hey are more universal and less exacting with respect to an optimization problem*≫ (see [12, page 9]).

Of course, PSO is not the only bio-inspired metaheuristics able to deal with minimization problems like ours. As possible alternatives we recall Differential Evolution (DE) and Genetic Algorithms (GAs). From a methodological point of view observe that our combined use of an exact penalty scheme with the PSO is not frequent in the literature. Indeed, for the solution of constrained problems the PSO is often modified with hybrid variants, which are suitably adapted to cope even with nonlinear constraints. In this respect we first provide a theoretical result ensuring the correspondence between the solutions of the original mathematical programming problem and the solutions of the exact penalty-based model (see Section 3 for details). Further, we also develop a simple original approach for the initialization of the particles' parameters. Its use yields numerical evidence of improvements in the convergence to a global minimum (see Section 4 for details). Finally, notice that the solution approach we propose is independent of the characteristics of the objective function and of the constraints. In other terms, our proposal might play the role of universal global (approximate) optimizer for a large variety of portfolio selection problems, even if characterized by very different risk measures and systems of constraints.

The remainder of this paper is organized as follows. In the next section first we illustrate the coherent risk measures we use, then we present our portfolio selection problem. In Section 3 first we recall the basics of PSO, then we apply an exact penalty method for the reformulation of our portfolio selection problem. In Section 4 we apply our overall solution procedure to the selection of large and complex portfolios, based on the set of assets constituting the Standard & Poor's SP 100 index. We test various settings of the solution procedure, where we use a simple approach for the initialization of the particles' parameters. Then, we apply our approach considering different time periods from August 2004 to October 2009, in order to detect possible differences in the optimal portfolio composition. Further, we compare these results with the ones coming from a couple of suitably chosen benchmark portfolio selection models. Final remarks are given in the last section.

---

[1] These boundings are known as cardinality constraints.

2. Portfolio selection and risk measures

The basic idea in the portfolio selection problem is to select stocks both maximizing portfolio performance and minimizing 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 using the expected value of the future return of the portfolio, the discussion regarding an adequate measure of risk is still open.

In its pioneering work [21] Markowitz proposed to use the variance of portfolio return to measure its risk, and this idea has been used for a long time in financial practice. However, it is well known that the mean–variance model leads to optimal investment decisions only if investment returns are elliptically distributed, or alternatively if the utility function of investors is quadratic. The main shortcomings of quadratic utility functions have been pointed out since their introduction (see [29]), and it is a stylized fact that the distributions of returns of financial instruments present asymmetry and "fat tails". These considerations have opened the way for the research on alternative measures of risk, along with their properties: a recent characterization of them is presented in [25].

One crucial fact should be taken into account for a correct specification of the risk measure: while variance gives the same weight to positive and negative deviations from the mean, several empirical studies have shown (see e.g. [18]) that investors treat them in different ways. This has led to the definition of risk measures that are focused on the "bad tail" of the distribution of the returns, as for example the semivariance (see [2]), the lower partial moments (see [14]) and the minimax ones (see [30]). On the other hand, some other risk measures were based on a quantile of the "bad tail", as the well known Value-at-Risk (VaR) (see [22]).

Since the introduction of the notion of coherent risk measure in [1], along with the specification of the properties for a measure (monotonicity, positive homogeneity, translation invariance and sub-additivity), there has been a growing interest for the previously introduced measures. In particular, properties similar to those for coherent risk measures have been studied also for the other risk measures. The Conditional Value-at-Risk (CVaR) (see [26]) is possibly the most famous measure obtained in this direction; other examples based on lower partial moments are reported in [13].

More recently Chen and Wang in [7] 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 the 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. Then, according with the conclusions presented in [7], 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 the class of risk measures in [7] for our portfolio selection problem. Our problem also takes into account several constraints, often used in fund management practice. In particular, we focus on handling the cardinality constraints, which yield a final model in the class of nonlinear mixed-integer programming problems.

2.1. Our portfolio selection model

Let Y be a real valued random variable defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and let us denote $\|Y\|_p = (E[|Y|^p])^{1/p}$, with $p \in [1, +\infty[$, where $E[\cdot]$ indicates the expected value of a random variable. Then, the measures of risk introduced in [7] are defined as:

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

where $a \in [0, 1]$, $Y^- = \max\{-Y, 0\}$ and $Y^+ = (-Y)^-$. For given a and p, any risk measure of this class is then a convex combination of the two coherent risk measures based on lower partial moments $\|(Y - E[Y])^-\|_1 - E[Y]$ and $\|(Y - E[Y])^-\|_p - E[Y]$. Thus, (1) is a coherent risk measure (see [13]). For a detailed description of its properties we refer the reader to [7]. 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.

Now we describe the portfolio selection model we consider. Suppose we have N assets to choose from, and for $i = 1, \ldots, N$ let $x_i \in \mathbb{R}$ be the weight of the i-th asset in the portfolio, with $X^T = (x_1, \ldots, x_N)$. Let $Z^T = (z_1, \ldots, z_N) \in \{0, 1\}^N$ be a binary vector, such that $z_i = 1$ if the i-th asset is included in the portfolio, $z_i = 0$ otherwise. Moreover, for $i = 1, \ldots, N$, 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. $\hat{r}_i = E[r_i]$. Then, the random variable $R \in \mathbb{R}$ 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, considering (1), our goal is to minimize $\rho_{a,p}(R)$, subject to several constraints. Of course the first ones to consider are the constraints regarding the minimum desirable expected return of the portfolio, i.e.

$$\hat{R} \geqslant l, \quad \text{with } l > 0,$$

along with the usual budget constraint

$$\sum_{i=1}^{N} x_i = 1.$$

Moreover, as stated in the previous section, we also introduce the following cardinality constraint: we select a (not too) small subset of the available assets. 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. When the number of selected assets is too large, several practical accounting problems may arise, which can increase transaction costs. By using the latter cardinality constraint we implicitly consider transaction costs in our model. The resulting constraint is explicitly given by

$$K_d \leqslant \sum_{i=1}^{N} z_i \leqslant K_u, \quad \text{where } 1 \leqslant K_d \leqslant K_u \leqslant N.$$

Further, we require that each of the selected assets cannot be a too large or too small fraction of the portfolio, i.e.

$$z_i d \leqslant x_i \leqslant z_i u, \quad \text{where } 0 \leqslant d \leqslant u \leqslant 1$$

and d, u represent respectively the minimum and maximum fraction allowed. Of course, to ensure compatibility with the cardinality constraint, the parameters d and u must satisfy

$$d \leqslant \frac{1}{K_d} \text{ and } u \geqslant \frac{1}{K_u}. \tag{2}$$

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

$$
\begin{aligned}
&\min_{X,Z} \quad \rho_{a,p}(R) \\
&\text{s.t. } \hat{R} \geqslant l \\
&\sum_{i=1}^{N} x_i = 1 \\
&K_d \leqslant \sum_{i=1}^{N} z_i \leqslant K_u \\
&z_i d \leqslant x_i \leqslant z_i u, \text{ with } i = 1, \ldots, N, \\
&z_i(z_i - 1) = 0, \text{ with } i = 1, \ldots, N,
\end{aligned}
\tag{3}
$$

where the last N constraints are introduced to model the relations $z_i \in \{0, 1\}$, with $i = 1, \ldots, N$. It is clear that if in the last N constraints we have $z_i = 0$, then the variable x_i does not play any role in the solution of problem (3), i.e. $x_i = 0$. Conversely, $z_i = 1$ implies that potentially the i-th asset will contribute to the final portfolio, with $x_i \in [d, u]$. Finally, the constraints $z_i(z_i - 1) = 0$, with $i = 1, \ldots, N$, represent just one (and possibly not the best) reformulation of the integrality constraints $z_i \in \{0, 1\}$, with $i = 1, \ldots, N$. We do not investigate further the latter issue, since it is not a focus of this paper.

Of course, (3) is a nonlinear and nonconvex mixed-integer problem, which in general admits several local solutions, but we want to possibly seek global solutions and not simply local minimizers. However, detecting precise solutions of (3) may be heavily time consuming in case exact methods are adopted. Thus, at present we experience the metaheuristic technique PSO on a non-smooth reformulation of problem (3). The next section is devoted to detail the PSO heuristics.

3. PSO for non-smooth reformulation of the portfolio selection problem

Particle Swarm Optimization is an iterative metaheuristics for the solution of nonlinear global optimization problems (see [17]). It is based on a biological paradigm, which is inspired by the flight of birds in a flock. In particular, the basic idea of PSO (see also [3] for a tutorial) is to replicate the behaviour of shoals of fishes or flocks of birds, when they cooperate in the search for food. On this purpose every member of the swarm explores the search area keeping memory of its best position reached so far, and it exchanges this information with the neighbors in the swarm. Thus, the whole swarm is supposed to converge eventually to the best global position reached by the swarm members.

In its mathematical counterpart the paradigm of a flying flock may be formulated as follows: given a minimization problem, find a global minimum (best global position) in a nonlinear 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 a random velocity, which is used to determine its initial direction of movement.

For a more formal description of PSO let us consider the global optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x}),$$

where $f : \mathbb{R}^d \mapsto \mathbb{R}$ is the objective function in the minimization problem. Suppose we apply PSO for its solution, where M particles are considered. At the k-th step of the PSO algorithm three vectors are associated to the j-th particle, $j \in \{1, \ldots, M\}$:

- $\mathbf{x}_j^k \in \mathbb{R}^d$, the position at step k of the j-th particle;
- $\mathbf{v}_j^k \in \mathbb{R}^d$, the velocity at step k of the j-th particle;
- $\mathbf{p}_j \in \mathbb{R}^d$, the best position visited so far by the j-th particle.

Moreover, $pbest_j = f(\mathbf{p}_j)$ denotes the value of the objective function in the position $\mathbf{p}_j$ of the j-th particle. The overall PSO algorithm, as in the version with inertia weight proposed in [27], is reported here:

1. Set $k = 1$ and evaluate $f(\mathbf{x}_j^k)$ for $j = 1, \ldots, M$. Set $pbest_j = +\infty$ for $j = 1, \ldots, M$.
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, with $j = 1, \ldots, M$, as

$$\mathbf{v}_j^{k+1} = w^{k+1} \mathbf{v}_j^k + \mathbf{U}_{\phi_1} \otimes (\mathbf{p}_j - \mathbf{x}_j^k) + \mathbf{U}_{\phi_2} \otimes (\mathbf{p}_{g(j)} - \mathbf{x}_j^k) \tag{4}$$

$$\mathbf{x}_j^{k+1} = \mathbf{x}_j^k + \mathbf{v}_j^{k+1} \tag{5}$$

where $\mathbf{U}_{\phi_1}, \mathbf{U}_{\phi_2} \in \mathbb{R}^d$ and their components are uniformly randomly distributed in $[0, \phi_1]$ and $[0, \phi_2]$ respectively. The symbol $\otimes$ denotes component-wise product and $\mathbf{p}_{g(j)}$ is the best position in a neighborhood of the j-th particle.
4. If a convergence test is not satisfied then set $k = k + 1$ and go to 2.

The values of ϕ_1 and ϕ_2 strongly affect the strength of the attractive forces, towards the personal and the neighborhood best positions explored so far by the j-th particle. Thus, in order to (possibly) yield the convergence of the swarm, they have to be set carefully in accordance with the value of the inertia weight w^k. The parameter w^k is generally linearly decreasing with the number of steps, i.e.

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

where common values for w_{max} and w_{min} are respectively 0.9 and 0.4, while K is usually the maximum number of steps allowed.

Another widely adopted version of the PSO algorithm is the one with constriction coefficients (see [8]), where the updating velocity rule (4) is replaced by

$$\mathbf{v}_j^{k+1} = \chi \left[\mathbf{v}_j^k + \mathbf{U}_{\phi_1} \otimes (\mathbf{p}_j - \mathbf{x}_j^k) + \mathbf{U}_{\phi_2} \otimes (\mathbf{p}_{g(j)} - \mathbf{x}_j^k) \right], \tag{6}$$

with $\chi = \frac{2}{\phi - 2 + \sqrt{\phi^2 - 4\phi}}$, $\phi = \phi_1 + \phi_2$, and $\phi > 4$.

As stated before, we can think that at step k the j-th particle moves as subject to two attractive vectors: the direction towards its previous best position (namely $(\mathbf{p}_j - \mathbf{x}_j^k)$) and the direction towards the best position in a suitable subset of the swarm (namely $(\mathbf{p}_{g(j)} - \mathbf{x}_j^k)$). We recall that $g(j)$ denotes the index of the particle with the best position reached so far, 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 $j = 1, \ldots, M$, and g is the index of the best particle in the whole swarm. This choice implies that the whole swarm is used as the neighborhood of each particle.

We remark that the original formulation of PSO was conceived for unconstrained problems. Thus, in general using PSO formulae 4,5, when constraints are included in the formulation, is improper. Indeed, in the latter case the algorithm above 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 [3]) 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 the random positioning strategies proposed in [32], or introducing some external criteria to rearrange the components of the particles, as the ones specific for cardinality constraints proposed in [20,10,28]. However, in this paper we decided to use PSO coherently with its original formulation, that is as a tool for the solution of unconstrained optimization problems. The latter choice is mainly motivated by the necessity of avoiding both a possible misleading application of specific metaheuristics to handle nonlinear constraints, and the careless setting of *ad hoc* coefficients.

To this purpose, first we have reformulated our problem into an unconstrained one, using the nondifferentiable ℓ_1 penalty function method described in [31,16]. The latter approach is known in the literature of constrained optimization as *exact penalty method*, where the term exact refers to the correspondence between the minimizers of the original constrained problem and the minimizers of the unconstrained (penalized) one. Some applications presented in the literature in which PSO is applied to portfolio selection problems for minimizing penalty functions or seemingly penalty functions are given in [6,28] respectively. Nevertheless, unlike our solution approach, the solution method proposed in [6] is not based on exact penalty

functions, and it does not consider integer unknowns. Moreover, the rule used in that paper for updating the parameter values appears to be far too problem dependent. Conversely, the solution method proposed in [28] has only the appearance of a penalty approach, but to a large extent it is based on so-called Lagrangian methods, inasmuch as the entire vector of the Lagrangian multipliers associated to the constraints is estimated at each iterate. With regard to that, generally Lagrangian methods have been proven to be effective in practice, but they require an accurate estimation of the Lagrangian multipliers, which is unnecessary for penalty-based solution approaches. Anyway, unlike our solution method, both such contributions do not ensure the correspondence between the solution of the original constrained problem and the reformulated one (see Proposition 3.1 for details).

On the contrary, in our approach we reformulate the problem (3) as follows (with $N + 1$ equality constraints and $2N + 3$ inequality constraints), using the nondifferentiable penalty function

$$\min_{X,Z} P(X,Z;\varepsilon) \tag{7}$$

where

$$P(X,Z;\varepsilon) = \rho_{a,p}(R) + \frac{1}{\varepsilon}\left[\max\{0, l - \hat{R}\} + \left|\sum_{i=1}^{N} x_i - 1\right| + \max\left\{0, K_d - \sum_{i=1}^{N} z_i\right\} + \max\left\{0, \sum_{i=1}^{N} z_i - K_u\right\}\right.$$
$$\left. + \sum_{i=1}^{N}\max\{0, z_i d - x_i\} + \sum_{i=1}^{N}\max\{0, x_i - z_i u\} + \sum_{i=1}^{N}|z_i(1 - z_i)|\right] \tag{8}$$

and ε is the penalty parameter.

The correct choice of ε ensures the correspondence between the solutions of problems (7) and (3) (see also [11]), which is summarized by the result which follows.

Proposition 3.1. *Consider problem* (3) *with* $\rho_{a,p}(R)$ *continuous on* $\mathbb{R}^N$. *Consider the Exact Penalty function* $P(X,Z;\varepsilon)$ *in* (8). *Let* (X^*, Z^*) *be a strict local minimizer of problem* (7) *where the KKT conditions (see the appendix for details) are satisfied, with the generalized Lagrange multipliers*

- λ_i^*, *with* $i = 1, \ldots, N + 1$ *(for the equality constraints),*
- σ_j^*, *with* $j = 1, \ldots, 2N + 3$ *(for the inequality constraints).*

Then, for any $\varepsilon \in (0, \varepsilon^*]$ *the solution* (X^*, Z^*) *is also a local minimizer of problem* (3), *where* ε^* *depends on* λ^* *and* σ^*.

Observe that the penalty function $P(X,Z;\varepsilon)$ is clearly nondifferentiable because of the ℓ_1-norm in (8). This also motivates the choice of using PSO for its minimization, since PSO evidently does not require the derivatives of $P(X,Z;\varepsilon)$. We avoid to go into details (see [11,31] for details), however the latter choice turns to be of great interest on those problems where illconditioning may arise. We also remark that the threshold value ε^* is unknown. Nevertheless, acceptable values of this threshold can often be found by appropriate numerical investigation, provided that a constraints qualification condition is satisfied in (X^*, Z^*).

Of course, since PSO is a metaheuristics, the minimization of the penalty function $P(X,Z;\varepsilon)$ theoretically does not ensure that a global minimum of the problem (3) is detected. Nevertheless, PSO often provides a suitable compromise between the performance (i.e. a satisfactory estimate of a global minimizer for the problem (3)) and the computational cost.

4. Numerical results

In this section, in order to describe the effectiveness of our approach, we focus on the following two issues:

- first we apply our solution procedure to the selection of portfolios based on the set of assets constituting the Standard & Poor's SP 100 index, over different time periods from August 2004 to October 2009;
- then we compare the so obtained results with the ones coming from a couple of benchmark portfolio selection models: a Markowitz-based solution procedure and a known alternative heuristic-based portfolio selection approach.

4.1. Applications and discussion

In order to test our approach, in this section we consider data of daily close prices $\{p_{i,t}\}$ of the i-th asset at time t, for $i = 1, \ldots, 100$ and $t = 1, \ldots, T$, where T is the time horizon considered. The assets considered are included in the Standard & Poor's SP 100 index,[2] from August 2004 to October 2009. Several subsets of these data have then been selected, to analyze

---

[2] Due to the lack of available data for the whole period considered, Mastercard and Philip Morris have been replaced by Verisign and Molson Coors Brewing Company.

differences in the optimal portfolio composition with respect to the considered time period. The price series have been used to compute each stock return

$$r_{i,t} = \frac{p_{i,t+1} - p_{i,t}}{p_{i,t}}.$$

Using the same idea of [7] we estimate the risk measure for any portfolio $X = (x_1, \ldots, x_N)^T$ as

$$\rho_{a,p}(R) = \frac{a}{T}\left[\sum_{t=1}^{T}\left(\sum_{i=1}^{N}(r_{i,t} - \hat{r}_i)x_i\right)^{+}\right] + (1-a)\left\{\frac{1}{T}\sum_{t=1}^{T}\left[\left(\sum_{i=1}^{N}(r_{i,t} - \hat{r}_i)x_i\right)^{-}\right]^{p}\right\}^{\frac{1}{p}}, \tag{9}$$

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

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

According with problem (3), the minimum level of desired return l in (8) was set to the global average of stock returns, i.e. $l = \sum_{i=1}^{N}\hat{r}_i$. Moreover, to reflect a realistic problem of portfolio selection, we set the values $d = 0.02$ and $u = 0.20$ in (8). In order to analyze the impact of meaningful cardinality constraints on the selection problem, we have chosen $K_d = 5$, while we have considered two different scenarios for $K_u : K_u = 50$ (the maximum value allowed according to (2)), and $K_u = 30$. The PSO algorithm to solve (8) was implemented in MATLAB 7, and the experiments have been performed on a workstation Acer Aspire M1610 with a Intel Core 2 Duo E4500 processor.

We first performed a set of preliminary tests (reported below) with the aim of assessing a proper setting for PSO. We experienced different values for the penalty parameter ε and for the coefficients of PSO, including the number M of particles in the swarm. Then, we chose the values of the parameters[3] a, p in the risk measure (9), as $a = 0.5$ and $p = 2$, and we ran the two versions of the algorithm (i.e. with decreasing inertia weight and constriction coefficients respectively), using one-year data of returns. Since the evaluation of the objective function $P(X, Z; \varepsilon)$ in (8) is relatively inexpensive, we stopped PSO iterations when either of the following stopping criteria was satisfied:

(a) the maximum number of 10 000 steps was outreached;
(b) $| fbest^{k+1} - fbest^k | < 10^{-8}$ for 2000 consecutive steps, where $fbest^k$ is the best value of the fitness function $f = P(X, Z; \varepsilon)$ at k-th iteration.

We remark that on average the version of PSO with constriction coefficients (6) showed earlier convergence to a global best position with worse values of the fitness function, compared to versions of PSO using a decreasing inertia weight. Thus, we decided to adopt the latter variant of PSO for the subsequent experiments.

In Table 1 we report the results in terms of the averaged best value of the fitness function (normalized to take into account the effects of using different values of ε) and its standard deviation. In Table 2 we show the results in terms of the averaged ratio between the final (F_f) and initial (F_i) value of the fitness function, including the average computational time, for different numbers of particles used. The results in Tables 1 and 2 are averaged over 10 runs of PSO, and Table 2 refers to the case $\varepsilon = 10^{-6}$, which is the best value from Table 1.

Table 2 also suggested to select $\varepsilon = 10^{-6}$ and $M = 200$. The number of particles is then quite high, but this concurs with the general evidence that larger populations perform better on higher dimensional problems (see also [4,5,3]).

Then, we repeated the computation in order to select the best values for the acceleration coefficients ϕ_1, ϕ_2, while for the initial and final values of the inertia weight w_{max}, w_{min} in (4) we used 0.9 and 0.4, as suggested by the current literature. The results are shown in Table 3, and the best performance was obtained with $\phi_1 = \phi_2 = 1.85$.

After this preliminary phase, we solved the portfolio selection problems for different values of the parameters a and p of the risk measure $\rho_{a,p}(R)$, and K_u, considering one year data of daily returns of different time periods. We wanted in fact to study both the capability of PSO to find a global minimum for the optimization problem and the economic effectiveness of the portfolios obtained, while considering different scenarios and different attitudes towards risk.

We then set the maximum number of the algorithm steps to 20 000 and, for every combinations of the parameters and the dataset, we first performed 25 runs of the algorithm, each with different random initial positions and velocities. The standard deviation of the values of the fitness function was still high (an example of the values of the fitness function at the end of the first 25 runs is shown in Table 4). Moreover, we found significantly different optimal portfolios in the runs, each corresponding to a possible local minimum. Then, we decided to iterate the procedure in the following way: we made another set of 25 PSO runs, with again random initial velocities for all particles. However, we used the 25 global best positions found in the previous phase as initial positions for 25 particles, while the remaining 175 ones were set randomly. At the end of this second phase we obtained convergence to the same global best position for each of the 25 runs (in general not corresponding to the best position of the previous 25 ones) and we assumed the latter to be the global minimum (X^*, Z^*) of the optimization

---

[3] Experiments with different values for a and p showed no significant difference with regard to the results of this preliminary phase.

Table 1
Results for different choices of the parameter ε in (8).

ε	Normalized fitness	Standard deviation
1	0.388728255	0.134295545
0.1	0.332573168	0.136412572
0.01	0.337188413	0.061398093
0.001	0.372167277	0.253545145
0.0001	0.413884094	0.155254628
0.00001	0.381672803	0.200184247
0.000001	0.260743870	0.099544544
0.0000001	0.341627127	0.185689216

Table 2
Results for different choices of the number M of particles in PSO.

M	Fitness	Ratio of decrease (F_f/F_i)	Time (seconds)
50	444393.5727	0.015960233	54.2939
100	60160.8053	0.002162255	109.2537
200	6423.0849	0.000233782	158.0136

Table 3
Results with different choices of the parameters ϕ_1 and ϕ_2.

$\phi_1 = \phi_2 = \phi/2$	Average fitness	Standard deviation
1.25	564611.56	261486.21
1.50	13439.35	11620.83
1.75	6047.39	5361.74
1.85	1822.09	1963.27
2.00	22763505.40	859690.31

Table 4
Best fitness after the first 25 runs, with $a = 0.5, p = 1, K_u = 50$, data of the period 2006–07.

Run	1	2	3	4	5
Best fitness	90.35	249.29	16.25	96.25	311.97
Run	6	7	8	9	10
Best fitness	247.11	4.96	57.84	50.78	384.16
Run	11	12	13	14	15
Best fitness	46.22	4.06	0.62	1.25	0.47
Run	16	17	18	19	20
Best fitness	2.17	0.22	112.70	118.46	0.29
Run	21	22	23	24	25
Best fitness	3.67	3.75	2.09	186.67	0.54

problem. We remark that the difference between the global best fitness $P(X^*, Z^*; \varepsilon)$ and the risk measure $\rho_{a,p}(R^*)$ was negligible,[4] except for the case $\rho_{0.5,5}$ with $K_u = 30$ and data of period 2006–07 (see Table 5). This means that the use of the nondifferentiable penalty function is effective, in order to impose the satisfaction of the constraints, including the cardinality constraints.

From Table 5 it is interesting to remark that the monotonicity properties expected by theoretical results ([7, Theorem 2.3], i.e. the monotonicity of $\rho_{a,p}(R)$ with respect to a and p) are fulfilled. Tables 6 and 7 also confirm the latter statement, and highlight that PSO is effective to detect a good approximation of a global minimum for $P(X, Z; \varepsilon)$. In this regard PSO provides an efficient compromise between the correctness of the solution found and the resources used in the computation.

Interestingly enough, we also notice that $\rho_{a,p}(R^*)$ slightly decreases with K_u (except for the case $\rho_{0.75,2}$ with data of period 2004–05). This is consistent with the fact that PSO has performed a good exploration of the feasible set when $K_u = 30$, that is a subset of the one with $K_u = 50$.

---

[4] At least in the cases where l is set equal to the global average of all stock returns, while there may be some problems when l is increased, as expected. We guess that in the latter case the maximum number of stocks to trade, i.e. K_u, is not enough large to fulfill the minimum return constraint.

Table 5
Comparison between the best fitness in the first 25 runs, and the global best fitness after other 25 runs, with $a = 0.5$, data of period 2006–07.

$K_u = 50$	$p = 1$	$p = 2$	$p = 5$
Best fitness among the first 25 runs	0.2208	0.6356	0.0662
Global best fitness after 25 + 25 runs	0.0020	0.0029	0.0044

$K_u = 30$	$p = 1$	$p = 2$	$p = 5$
Best fitness among the first 25 runs	1.4742	0.4846	0.2062
Global best fitness after 25 + 25 runs	0.0021	0.0031	0.0134

Table 6
Monotonicity of $\rho_{a,p}(R^*)$ for $a = 0.5$ and different values of p and K_u, with one year data from two time periods. $N_{a,p}(K_u)$ represents the number of assets in the final portfolio.

2004–05	$p = 1$	$p = 2$	$p = 5$
$\rho_{0.5,p}; K_u = 50$	0.001951	0.002816	0.004513
$N_{a,p}(K_u)$	43	45	40
$\rho_{0.5,p}; K_u = 30$	0.002333	0.003219	0.004808
$N_{a,p}(K_u)$	23	28	28

2006–07	$p = 1$	$p = 2$	$p = 5$
$\rho_{0.5,p}; K_u = 50$	0.002012	0.002924	0.004379
$N_{a,p}(K_u)$	44	45	45
$\rho_{0.5,p}; K_u = 30$	0.002094	0.003099	0.004339
$N_{a,p}(K_u)$	30	30	29

Table 7
Monotonicity of $\rho_{a,p}(R^*)$ for $p = 2$ and different values of a and K_u, with one year data from two time periods. $N_{a,p}(K_u)$ represents the number of assets in the final portfolio.

2004–05	$a = 0$	$a = 0.25$	$a = 0.5$	$a = 0.75$	$a = 1$
$\rho_{a,2}; K_u = 50$	0.003853	0.003448	0.002816	0.002509	0.002135
$N_{a,p}(K_u)$	43	45	45	46	43
$\rho_{a,2}; K_u = 30$	0.004029	0.003473	0.003219	0.002434	0.002306
$N_{a,p}(K_u)$	29	29	28	27	29

2006–07	$a = 0$	$a = 0.25$	$a = 0.5$	$a = 0.75$	$a = 1$
$\rho_{a,2}; K_u = 50$	0.004194	0.003587	0.002924	0.002319	0.001820
$N_{a,p}(K_u)$	45	41	45	48	48
$\rho_{a,2}; K_u = 30$	0.004227	0.003652	0.003099	0.002345	0.001855
$N_{a,p}(K_u)$	30	30	30	29	29

4.2. Comparisons and discussion

In this section we compare the performances of the obtained portfolios with the ones of some other portfolios, inspired by practical rules adopted by fund management industry.

4.2.1. The Markowitz-based benchmark

First we followed some indications of the previously mentioned north-eastern Italian company, which is specialized in automatic financial trading systems. According with their policy, several professional investors often tend to select their portfolios still using simple Markowitz-like approaches. Thus, we first compared our portfolios with the ones obtained by solving a new selection problem, in which we both replaced $\rho_{a,p}(R)$ with variance (the quintessential risk measure à la Markowitz) and kept the same system of constraints of problem (3). The optimization problem associated to this new portfolio selection problem was again solved by using PSO.

Then, in a second experience we compared our portfolio with the one obtained by the very basic Markowitz portfolio selection model (see below for details). In this way we performed a comparison with a widely-if-not-universally accepted benchmark in portfolio management.

According with the plan of experiments above, we first checked for the diversification of the portfolios in the different approaches, as shown in Table 8. It appears that when $K_u = 50$ (i.e. the cardinality constraint is more relaxed), the diversification obtained using $\rho_{a,p}(R)$ is higher than using variance, and it is also slightly increasing with p. Again, this is also consistent with the results obtained in [7], where the cardinality constraint was not explicitly introduced, and the comparison was carried on with respect to CVaR as risk measure.

Table 8
Comparison between the number of assets $N_{a,p}(K_u)$ in the optimal portfolios, using respectively $\rho_{0.5,p}(R)$ and the variance σ^2 as the objective function in (3), with one year data from two time periods.

2004–05	$\rho_{0.5,p}$			σ^2
	$p = 1$	$p = 2$	$p = 5$	
$K_u = 50$	43	45	45	42
$K_u = 30$	29	28	28	28

2006–07	$\rho_{0.5,p}$			σ^2
	$p = 1$	$p = 2$	$p = 5$	
$K_u = 50$	44	45	45	42
$K_u = 30$	30	30	29	28

Then, we compared the performance of the portfolios selected using respectively $\rho_{a,p}(R)$ and σ^2 in (3). Following indications from financial practice, we proceeded as follows: we used one-year data of daily returns, from the portfolio selected by minimizing $P(X,Z;\varepsilon)$ in (8). Then, we invested the selected portfolios for the next three months. After that we repeated the selection, and we re-invested the resulting portfolios for other three months, and so on for other two quarters. In this way we analyzed the performance of the portfolios along one entire year. We considered two periods of one year length (in order to compute the portfolios for the test in the subsequent first quarter), that is August 2004-July 2005 and February 2007-January 2008, with the aim of analyzing the impact of different macroeconomic conditions on the performance of the portfolios. The results are shown in Tables 9–12.

We notice that the behaviour of the portfolio selected using $\rho_{0.5,2}$ is quite similar to the one of the portfolio selected using the variance. The latter fact is more evident when $K_u = 30$ in the period 2005–06, where in two cases we obtain the same

Table 9
Portfolio returns with different measures of risk (namely $\rho_{0.5,p}(R)$ and σ^2 in (8)); case with $K_u = 50$ for the time period 2005–06.

Period length	$\rho_{0.5,1}$ (%)	$\rho_{0.5,2}$ (%)	$\rho_{0.5,5}$ (%)	σ^2 (%)
3 months	1.67	1.33	0.67	−0.18
6 months	3.22	4.00	4.55	3.59
9 months	0.47	−1.41	1.14	3.35
12 months	−3.35	−2.79	−1.24	−2.59

Table 10
Portfolio returns with different measures of risk (namely $\rho_{0.5,p}(R)$ and σ^2 in (8)); case with $K_u = 30$ for the time period 2005–06.

Period length	$\rho_{0.5,1}$ (%)	$\rho_{0.5,2}$ (%)	$\rho_{0.5,5}$ (%)	σ^2 (%)
3 months	−1.54	1.68	0.54	3.17
6 months	4.68	5.36	3.05	5.36
9 months	2.13	0.96	1.14	0.96
12 months	−3.23	−1.89	−0.93	−1.89

Table 11
Portfolio returns with different measures of risk (namely $\rho_{0.5,p}(R)$ and σ^2 in (8)); case with $K_u = 50$ for the time period 2008–09.

Period length	$\rho_{0.5,1}$ (%)	$\rho_{0.5,2}$ (%)	$\rho_{0.5,5}$ (%)	σ^2 (%)
3 months	7.23	3.20	3.19	5.13
6 months	−8.32	−6.64	−6.79	−8.24
9 months	−23.82	−25.13	−22.81	−22.30
12 months	−10.44	−9.77	−9.38	−7.11

Table 12
Portfolio returns with different measures of risk (namely $\rho_{0.5,p}(R)$ and σ^2 in (8)); case with $K_u = 30$ for the time period 2008–09.

Period length	$\rho_{0.5,1}$ (%)	$\rho_{0.5,2}$ (%)	$\rho_{0.5,5}$ (%)	σ^2 (%)
3 months	3.42	3.63	3.94	3.53
6 months	−8.10	−9.18	−6.89	−10.24
9 months	−30.50	−24.42	−19.84	−19.52
12 months	−13.15	−11.85	−10.51	−12.10

Table 13
Portfolio returns for the basic Markowitz model (i.e. $\rho_{0.5,p}(R)$ in (3) is replaced by σ^2, and the feasible set of (3) contains only the first two constraints).

Period length	2005–06 (%)	2008–09 (%)
3 months	−3.39	−1.27
6 months	3.32	−5.73
9 months	3.35	−9.18
12 months	−2.32	−12.25

return and the differences in the composition of the two portfolios are negligible. On the contrary, the portfolios selected using $\rho_{0.5,1}$ and $\rho_{0.5,5}$ appear to correspond respectively to a more aggressive and conservative investor. In particular, during the financial markets crisis in the period 2008–09, the portfolio selected using $\rho_{0.5,5}$ shows less losses than the ones obtained using variance. These results are definitely expected, considering the theoretical role of the parameter p described in [7].

As final comparison we report in Table 13 the returns of portfolios obtained using the very basic Markowitz model (i.e. the returns of portfolios selected using variance as measure of risk, and keeping only the first two constraints of problem (3)), which is still very used in the financial practice with minimal modifications. We remark that in this case there are no restrictions on either the number of assets or the fraction of the portfolio invested in each asset; thus, all assets are possibly involved in the composition of the optimal portfolio. This in practice entails a notable increase of the related transaction costs. Moreover, observing the returns on the "good" period 2005–06, we notice that the higher diversification provided by this simple model does not yield better performance with respect to the portfolios selected using $\rho_{a,p}(R)$ or variance in problem (3). In the "bad" period 2008–09 the performance of Markowitz portfolios is slightly better. This is not unexpected, since in this case short sales are allowed, and during this time period almost all assets show negative returns. On the other hand, short positions generally require higher transaction costs, including interest expenses for borrowing assets, and they could be not allowed by markets supervisor during big financial crisis. Thus, to sum up portfolios selected using the general model (3) could still represent a better alternative, especially for a fund manager.

4.2.2. The alternative heuristic-based benchmark

Finally, we compare the results obtained by our solution procedure with the ones coming from a known heuristic-based approach. As we premised in Section 1, the cardinality constraints play a remarkable role in our portfolio selection model. Indeed, on one hand they are often considered by the fund manager industry, moreover they make the corresponding mathematical programming problem NP-hard. Thus, in order to check the capability of PSO to handle such constraints, and to closely approach a global solution of problem (3), we used an heuristic-based procedure as a benchmark. The latter method is often used when facing cardinality constrained problems (see for example [19, page 117]). Below we synthetically describe this alternative heuristic-based approach. For each considered combination of the values for a and p in $\rho_{a,p}(R)$:

- first, we sort in an increasing manner the assets on the basis of their performances with respect to $\rho_{a,p}(R)$;
- then, we consider the set $\widetilde{X} \subseteq X$ constituted by the first $N_{a,p}(K_u)$ assets of the ordered list, where $N_{a,p}(K_u)$ is the number of assets in the optimal portfolio reported in Tables 6 and 7, and determined by the PSO-based approach in the full version of the problem;
- finally, to assign the weights to the assets we solved the following reduced optimization problem:

$$\min_{\widetilde{X}} \quad \rho_{a,p}(R)$$
$$\text{s.t. } \hat{R} \geqslant l$$
$$\sum_{i=1}^{N_{a,p}(K_u)} \widetilde{x}_i = 1 \tag{10}$$
$$d \leqslant \widetilde{x}_i \leqslant u, \text{ with } i = 1, \ldots, N_{a,p}(K_u).$$

Notice that:

- we have assumed that the number of assets selected both in the full version of the problem and in the reduced one has to be the same, otherwise the portfolio with a higher number of selected assets should have had more potential advantages in terms of diversification;
- the cardinality constraint is "managed" externally to the reduced optimization problem.

For comparability reasons, we decided again to reformulate problem (10) in a non-smooth way using the nondifferentiable ℓ_1 penalty function method, similarly to what we have done for problem (3); then, we used again PSO for its solution. The results, in term of the global risk measure of the portfolio $\rho_{a,p}(R)$ are reported in Table 14 and show, by comparison with Table 6, how the best values are generally obtained by PSO in the solution of the full problem (3). In particular, the percent-

Table 14
Values of $\rho_{a,p}(R^*)$ of problem (10) for $a = 0.5$ and different values of p and $N_{a,p}(K_u)$, with one year data from two time periods.

2004–05	$p = 1$	$p = 2$	$p = 5$
$\rho_{0.5,p}; N_{a,p}(K_u = 50)$	0.002143	0.003261	0.004592
	9.84%	15.80%	1.75%
$\rho_{0.5,p}; N_{a,p}(K_u = 30)$	0.002437	0.003518	0.004796
	4.46%	9.29%	−0.25%
2006–07	$p = 1$	$p = 2$	$p = 5$
$\rho_{0.5,p}; N_{a,p}(K_u = 50)$	0.002452	0.003126	0.004459
	21.87%	6.91%	1.83%
$\rho_{0.5,p}; N_{a,p}(K_u = 30)$	0.002293	0.003379	0.004401
	9.50%	9.04%	1.43%

ages reported in Table 14, under each best value, measure the variation of such values with respect to the corresponding values associated to the full problem (3). All but one of these percentages are positive and not particularly close to 0 (the only one negative percentage is pretty close to 0). This shows a general capability of our solution approach (i.e. the full model (3) with the nonsmooth reformulation (8)) to fruitfully handle the cardinality constraints better than the considered benchmark portfolio selection alternative. We highlight that such a capability is reasonably more evident in the case when $K_u = 50$, that is the case in which there is more diversification (which is well exploited by our PSO-based algorithm).

5. Final remarks

In this paper we have first proposed a partially novel reformulation for the selection of large and complex portfolios, characterized by an upper-and-lower-moments-based coherent risk measure and a mixed-integer formulation. This reformulation used a nondifferentiable exact penalty method and was solved using PSO. Although the obtained results are satisfactory, this solution approach seems to offer opportunities for possible improvements and extensions. In particular:

- As the reformulation of problem (3) is concerned, we point out that other possible reformulations of that mathematical programming problem may be considered, both smooth and non-smooth. To this purpose, resorting to continuously differentiable penalty functions appears particularly promising. This method is substantially characterized by theoretical properties equivalent to the ones characterizing the penalty scheme used in this paper (see [15] for details);
- As the initialization of the particles' positions and velocities is concerned, we guess that the performances of our simple approach can be significantly improved by resorting to a theoretical-based procedure recently proposed in [4,5]. By so doing we expect improvements both in the solution quality and computational time;
- As stated in Section 1, PSO in not the only bio-inspired metaheuristic able to deal with minimization problems like (8). Currently, in order to compare different bio-inspired metaheuristics as global minimizers of complex portfolio selection problems, we have started to use GAs. The very first preliminary results (not included here) suggest that the optimal portfolio compositions obtained by PSO and GAs are quite similar. However, PSO needs a computational time which is significantly lower than the one needed by GAs;
- Finally we recall that, from a methodological point of view, the solution approach we propose can play the role of universal global (approximate) optimizer for a large variety of complex portfolio selection problems, even if characterized by very different risk measures and systems of constraints. Thus, it can represent a flexible tool for the fund management industry, in order not to leave unsatisfied demand.

In order to carefully detect features and drawbacks of our approach, in a future research further investigations are necessary with respect to different risk measures, constraints and data.

Acknowledgements

Appendix A

Consider the general constrained optimization problem

$$\min \quad f(x)$$
$$h_i(x) = 0 \quad i = 1, \ldots, p,$$
$$g_j(x) \leqslant 0 \quad j = 1, \ldots, q. \tag{11}$$

Suppose that at the feasible point x^* some inequality constraints (thereof subscripts are in the subset $\mathcal{A}(x^*)$) are satisfied as equalities, i.e.

$$g_j(x^*) = 0, \quad j \in \mathcal{A}(x^*).$$

We say that for the problem (11) the condition *LICQ* (Linear Independent Constraint Qualification) holds at x^* if the vectors

$$\left\{ \nabla h_1(x^*), \ldots, \nabla h_p(x^*), \nabla g_j(x^*)\big|_{j \in \mathcal{A}(x^*)} \right\}$$

are linearly independent. Then we can now define the following first order optimality conditions for the minimizer x^* of (11).

Proposition 5.1 (*KKT Conditions*). *Consider the problem* (11), *where the functions f, h and g are continuously differentiable. Suppose that x^* is a local minimizer of* (11), *where the LICQ holds. Then, there exists a unique Lagrange multiplier vector $(\lambda^{*T}, \sigma^{*T}) \in \mathbb{R}^{p+q}$ such that*

$$\nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{q} \sigma_j^* \nabla g_j(x^*) = 0$$
$$h_i(x^*) = 0, \quad i = 1, \ldots, p$$
$$\sigma_j^* g_j(x^*) = 0, \quad j = 1, \ldots, q$$
$$g_j(x^*) \leqslant 0, \quad j = 1, \ldots, q$$
$$\sigma_j^* \geqslant 0, \quad j = 1, \ldots, q.$$

Observe that the constraints qualification condition *LICQ* in Proposition 5.1 substantially ensures that there exist the functions $\lambda = \lambda(x)$ and $\sigma = \sigma(x)$, with $\lambda^* = \lambda(x^*)$ and $\sigma^* = \sigma(x^*)$, which can be explicited by the Implicit Function Theorem, at least in a neighborhood of x^*. Equivalently, the condition *LICQ* can be replaced by several other qualification conditions (see also [16,11,31]).

References

[1] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent measures of risk, Mathematical Finance 9 (3) (1999) 203–228.
[2] E. Ballestero, Mean-semivariance efficient frontier: a downside risk model for portfolio selection, Applied Mathematical Finance 12 (1) (2005) 1–15.
[3] T. Blackwell, J. Kennedy, R. Poli, Particle swarm optimization – An overview, Swarm Intelligence 1 (1) (2007) 33–57.
[4] E.F. Campana, G. Fasano, A. Pinto, Dynamical analysis for the selection of parameters and initial population in Particle Swarm Optimization, Journal of Global Optimization 48 (3) (2010) 347–397.
[5] E.F. Campana, M. Diez, G. Fasano, D. Peri, Improving the initial particles position and parameters selection for PSO in bound constrained optimization problems, Springer Lecture Notes in Computer Science, vol. 7928, Springer, Berlin Heidelberg, 2013, pp. 112–119.
[6] W. Chen, R. Zhang, Y. Cai, F. Xu, Particle swarm optimization for constrained portfolio selection problems, in: Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 2006, pp. 2425–2429.
[7] Z. Chen, Y. Wang, Two-sided coherent risk measures and their application in realistic portfolio optimization, Journal of Banking and Finance 32 (12) (2008) 2667–2673.
[8] M. Clerc, J. Kennedy, The Particle swarm – Explosion, stability, and convergence in a multidimensional complex space, IEEE Transaction on Evolutionary Computation 6 (1) (2002) 58–73.
[9] A.R. Conn, Constrained optimization using a nondifferentiable penalty function, SIAM Journal on Numerical Analysis 10 (4) (1973) 760–784.
[10] T. Cura, Particle swarm optimization approach to portfolio optimization, Nonlinear Analysis: Real World Applications 10 (4) (2009) 2396–2406.
[11] G. Di Pillo, L. Grippo, Exact penalty functions in constrained optimization, SIAM Journal on Control and Optimization 27 (6) (1989) 1333–1360.
[12] V. Feoktistov, Differential Evolution: In Search of Solutions, Springer, New York, 2006.
[13] T. Fischer, Risk capital allocation by coherent risk measures based on one-sided moments, Insurance: Mathematics and Economics 32 (1) (2003) 135–146.
[14] P.C. Fishburn, Mean-risk analysis with risk associated with below-target returns, American Economic Review 67 (2) (1977) 116–126.
[15] R. Fletcher, A class of methods for nonlinear programming with termination and convergence properties, in: J. Abadie (Ed.), Integer and Nonlinear Programming, North-Holland, Amsterdam, 1970, pp. 157–175.
[16] R. Fletcher, Practical Methods of Optimization, John Wiley & Sons, Glichester, 1991.
[17] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in: Proceedings of the IEEE International Conference on Neural Networks, San Francisco, 1995, pp. 1942–1948.
[18] L. Knetsch, B.F.J. Borges, Tests of market outcomes with asymmetric valuations of gains and losses: smaller gains, fewer trades, and less value, Journal of Economic Behavior & Organization 33 (2) (1998) 185–193.
[19] D. Maringer, Portfolio Management with Heuristic Optimization, Springer, Dordrecht, 2005.
[20] D. Maringer, O. Oyewumi, Index tracking with constrained portfolios, Intelligent Systems in Accounting, Finance and Management 15 (1–2) (2007) 57–71.
[21] H.M. Markowitz, Portfolio selection, Journal of Finance 7 (1) (1952) 77–91.
[22] A.J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, Princeton, 2005.
[23] C.H. Papadimitriou, K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, New Jersey, 1982.
[24] T. Pietrzykowski, An exact potential method for constrained maxima, SIAM Journal on Numerical Analysis 6 (2) (1969) 299–304.
[25] S. Rachev, S. Ortobelli, S. Stoyanov, F.J. Fabozzi, A. Biglova, Desirable properties of an ideal risk measure in portfolio theory, International Journal of Theoretical and Applied Finance 11 (1) (2008) 19–54.
[26] R.T. Rockafellar, S. Uryasev, Conditional value-at-risk for general loss distributions, Journal of Banking and Finance 26 (7) (2002) 1443–1471.

[27] Y. Shi, R. Eberhart, A modified particle swarm optimizer, in: The 1998 IEEE International Conference on Evolutionary Computation Proceedings 1998. IEEE World Congress on Computational Intelligence, 1998, pp. 69–73.

[28] N. Thomaidis, T. Angelidis, V. Vassiliadis, G. Dounias, Active portfolio management with cardinality constraints: an application of particle swarm optimization, New Mathematics and Natural Computation 5 (3) (2009) 535–555.

[29] R.F. Wippern, Utility implications of portfolio selection and performance appraisal model, Journal of Financial and Quantitative Analysis 6 (3) (1971) 913–924.

[30] M.R. Young, A minimax portfolio selection rule with linear programming solution, Management Science 44 (5) (1998) 673–683.

[31] W.I. Zangwill, Non-linear programming via penalty functions, Management Science 13 (5) (1967) 344–358.

[32] W.J. Zhang, X.F. Xie, D.C. Bi, Handling boundary constraints for numerical optimization by particle swarm flying in periodic search space, arXiv:cs/ 0505069, 2005.