Empirical Analysis of Index Tracking Error Minimization Algorithms Based on Stochastic Dominance Principle*

Alexey R. Faizliev, Sergei P. Sidorov, Sergei V. Mironov, and Andrey A. Khomchenko

Saratov State University, Russia
faizlievar1983@mail.ru

Abstract. Index tracking strategy is a passive financial strategy aiming at replication of a given index or portfolio return. In this study, a solution for the problem of index tracking is regarded with account of cardinality constraint, i. e. with a restriction on the maximum amount of assets held in the portfolio. The article discusses different algorithms to solve this problem in l_2-norm, specifically, greedy algorithm, differential evolution algorithm, and LASSO-type algorithm. For the empirical analysis we used public data relating to the three major market indices: Hang Seng (Hong Kong), S&P 100 (USA) and Nikkei 225 (Japan). For comparative analysis of greedy algorithm with LASSO-type algorithm and differential evolution algorithm stochastic dominance principle was used. At that, the comparison of the approaches included in-sample data as well as out-of-sample data.

Keywords: index tracking; decision making, portfolio optimization, greedy algorithms, differential evolution algorithms, stochastic dominance

1 Introduction

For any $q > 0$ and $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, let $\|x\|_q := \left(\sum_{i=1}^n |x_i|^q\right)^{1/q}$ and $\|x\|_0 = \lim_{q\to 0+} \|x\|_q =$ (the number of non-zero elements of x). If $q \geq 1$ then $\|x\|_q$ denotes l_q-norm of $x \in \mathbb{R}^n$. Let n be the number of investable assets. Denote r_{ti} the return of asset i at time t, $1 \leq i \leq n$, $1 \leq t \leq m$, $R = (r_{ti})$ is the $m \times n$ matrix. A portfolio is defined to be a vector of weights, $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$. We do not allow the portfolio changes over time and do not take into account transaction costs. We will assume that

1. one unit of capital is available, i. e. $x^T 1_n = 1$, where 1_n denotes the vector from $\mathbb{R}^n$ in which every component is equal to 1;
2. short selling is allowed, i. e. weights x_i can be negative.

* The results were obtained within the framework of the state task of RFBR (grant 14-01-00140).

Let I_t be the index return at time t, $1 \le t \le m$, and $I = (I_1, \dots, I_t)^T \in \mathbb{R}^m$. In the traditional index tracking optimization, the objective is to find a portfolio which has minimal tracking error variance, the sum of squared deviations between portfolio returns and market index returns (see e. g. [20]):

$$x^* = \arg\min \frac{1}{m} \|I - Rx\|_2^2 \quad s.\,t. \quad x^T 1_n = 1. \tag{1}$$

It should be noted that the standard Markovitz model is a special case of index tracking portfolio model (1) (see, for example [4, 25]). Since the problem (1) is the problem of convex optimization, it can be easily solved by Lagrange method.

Instead of squared deviations, the absolute error is also presented in [9, 13, 14, 18, 19, 21, 23, 28] and is used in practice.

In this paper we will examine three algorithms for solving the problem (1) with the cardinality constraint:

$$x^* = \arg\min \frac{1}{m} \|I^T 1_n - Rx\|_2^2 \;\; \text{s. t.} \;\; x^T 1_n = 1, \; \|x\|_0 \le K, \tag{2}$$

where K is the limit on the number of assets in the portfolio with non-zero weights. It is supposed that K is substantially smaller than n, $K \ll n$.

Index tracking problem with cardinality constraint is NP-hard problem and it usually requires the development of heuristic algorithms such as genetic algorithms and differential evolution algorithm [1,8–12,16]. These algorithms, though providing a sufficiently accurate solution of the problem, are associated with significant time costs connected with algorithm operation. Good reviews can be found in [1, 5, 17]. Greedy algorithms also proved their effectiveness [7, 25]. On the other hand, greedy algorithms does not necessarily yield an optimal solution.

In section 2 we describe three algorithms for solving the problem (2):

- the greedy algorithm (GA),
- the LASSO-type algorithm (LASSO),
- the differential evolution algorithms (DE).

Therefore, our comparative analysis will be based on algorithms from three different classes of algorithms: GA is a trajectory method, DE-algorithm is a population-based method and LASSO-type algorithm is method with the cardinality constraint relaxation.

The greedy algorithm presented in this paper uses the adaptation of ideas of [25]. In section 3, using a technique for comparative analysis based on both first order stochastic dominance and second order stochastic dominance principles, we compare the performance of three different portfolios obtained by the three algorithms for index tracking problem (2).

2 Algorithms for Minimization of Index Tracking Error in l_2-norm with the Cardinality Constraint

2.1 Greedy algorithm for l_2-norm minimization with regularization

Let $N = \{1, \ldots, n\}$ be the index set of investable assets. Problem (2) is the special case (with $\tau = 0$) of the following problem

$$x^* = \arg\min_x \|I - Rx\|_2^2 + \tau\|x\|_2^2 \;\; s.t.\; x^T 1_n = 1,\; \|x\|_0 \le K, \tag{3}$$

where τ is a positive parameter. The regularization term $\tau\|x\|_2^2$ allows to use the least squares estimator even in the case of multicollinearity in the matrix R [25]. The model with $\tau = 0$ was examined in [6,27]. The greedy algorithm for solving the problem (3) in l_2-norm was studied in [25]. The algorithm at each step adds to our portfolio the asset that is the closest to the index. The process continues until we reach the cardinality K.

Let $M_k \subset N$ be the set of indices corresponding to k non-zero elements of x. Let $\widetilde{R}_{M_k}$ be a submatrix of R with dimension $(m \times |M_k|)$. Then the problem (3) with $x_i = 0$ for $i \in N \setminus M_k$ can be rewritten as

$$\widetilde{x}^* = \arg\min_{\widetilde{x}} \|I - \widetilde{R}_{M_k}\widetilde{x}\|_2^2 + \tau\|\widetilde{x}\|_2^2 \;\; s.t.\; \widetilde{x}^T 1_{|M_k|} = 1,\; \widetilde{x} \in \mathbb{R}^{|M_k|}. \tag{4}$$

Denote $f^\tau(M_k) := \|I - \widetilde{R}_{M_k}\widetilde{x}^*\|_2^2 + \tau\|\widetilde{x}^*\|_2^2$. The optimal solution of the problem (4) can be obtained by Lagrange method:

$$\widetilde{x}^\tau_{M_k} = (\widetilde{R}^T_{M_k}\widetilde{R}_{M_k} + \tau E_k)^{-1}(\widetilde{R}^T_{M_k} I - \lambda e_k), \tag{5}$$

where E_k is the $(k \times k)$-identity matrix and $\lambda = \frac{1_k^T(\widetilde{R}^T_{M_k}\widetilde{R}_{M_k} + \tau E_k)^{-1}\widetilde{R}^T_{M_k} I - 1}{1_k^T(\widetilde{R}^T_{M_k}\widetilde{R}_{M_k} + \tau E_k)^{-1} 1_k}$.

Algorithm 1: Greedy Algorithm in l_2

---

begin
- · Let $M_0 = \varnothing$ and $k = 1$. Set $f^\tau(M_0)$ to be sufficiently large;
- · **while** $k \le K$ **do**
  - · $\forall s \in N \setminus M_{k-1}$ calculate $\widetilde{x}^\tau_{M_{k-1}\cup\{s\}}$ using (5);
  - · Select $s^* = \arg\min_{s \in N\setminus M_{k-1}} f^\tau(M_{k-1} \cup \{s\})$ and set $\widetilde{x}^\tau_{M_k} = \widetilde{x}^\tau_{M_{k-1}\cup\{s^*\}}$;
  - · Set $M_k = M_{k-1} \cup \{s^*\}$ and $k = k+1$;
- · Set $x^\tau_G = \widetilde{x}^\tau_{M_K}$ and $M^\tau_G = M_K$;
- · **return** x^τ_G and M^τ_G;

end

2.2 LASSO-type Algorithm

Minimization of l_1 penalized objective functions can have a sparsifying effect that has long been observed in research and practice. Minimizing l_1 norm is now a widely used technique for obtaining sparse solutions [4]. The paper [4] uses LASSO-type approach when the problem (1) is reformulated as a constrained least-squares regression problem. Let us consider the problem

$$x^{\delta} = \arg\min \|x\|_1 \quad s.t. \quad \|I - Rx\|_2 \le \delta, \; x^T 1_n = 1, \tag{6}$$

where δ is a scalar, which is assumed to be selected so that the true vector is feasible with high probability. The problem of the type (6) without the constraint $x^T 1_n = 1$ is called LASSO regression [26]. LASSO-type estimator is generally able to accurately estimate nearly sparse vectors. Effective algorithms for sparse recovery applications problem were developed in the paper [3]. For numerical solution of the problem (6) we used the set of Matlab templates TFOCS, that accompany the paper [3].

The number of assets in the portfolio with non-zero weights (i. e. cardinality K) of the optimal solution of the problem (6) depends on parameter δ. Bigger (smaller) values of δ correspond to smaller (bigger) values of K.

2.3 Differential evolution algorithm for l_2-norm minimization

A recent addition to the class of heuristics is the evolutionary method of differential evolution (DE) proposed by [24].

In our work to solve the problem (2), we use the algorithm of differential evolution. DE algorithm is one of the possible "continuous" modifications of standard genetic algorithms. At the same time, this algorithm has one essential feature that largely determines its properties. The "inner" random number generator is used as a source of external noise, and implemented as the difference between random vectors of the selected population.

Let N be the number of portfolios in each population. The initial population P of portfolios $x^i = (x_0^i, \ldots, x_n^i)^T \in \mathbb{R}^n$, $i = 1, \ldots, N$, is obtained as follows. First, we randomly generate N vectors y^i from $\mathbb{R}^n$. For each y^i we set to zero $n - K$ elements y_j^i of vector y^i that are closest to zero (the cardinality of y^i becomes K) and then we set $x_j^i = y_j^i / (\sum_{s=1}^n y_s^i)$ to gain restriction $x^T 1_n = 1$.

Portfolios x^i refer to the point of n-dimensional space which defines the objective function $\frac{1}{m}\|I - Rx\|_2$ that we want to minimize. At each iteration, the algorithm produces a new generation of portfolios (population) randomly combined from portfolios of the previous generation. Portfolios of new generation is generated in a few steps. First, for each $i = 1, \ldots, N$ we randomly select three different portfolios x^a, x^b, x^c among the portfolios of the previous generation. Then, we calculate

$$\tilde{x}_j^i = x_j^a + (F + z_1)(x_j^b - x_j^c + z_2)$$

where $\tilde{x}_j^i$, x_j^a, x_j^b, x_j^c are j-th components of $\tilde{x}^i$, x^a, x^b, x^c vectors respectively. Parameter F is a positive real constant from the interval $[0, 2]$, which manages the increasing influence of the difference $x_j^b - x_j^c + z_2$ in the result vector, z_1 and z_2 are either equal to zero with small probabilities (for example, 0.001 and 0.002, respectively), or they are normally distributed random variables with mean zero, and a small standard deviation (i. e. 0.002). Then we set to zero $n - K$ elements of $\tilde{x}^i$ (the cardinality of $\tilde{x}^i$ must be equal to K), and after it we set $\tilde{x}_j^i = \hat{x}_j^i / (\sum_{s=0}^n \tilde{x}_s^i)$ to fulfil the budget constraint.

Parameters z_1 and z_2 are optional parameters of differential evolution algorithm; they make "noise" in the calculation of the resulting vector which helps to avoid falling into local extremes.

Component $\hat{x}_j^i$ of vector $\hat{x}^i$ replaces x_j^i with probability π and the portfolio $\hat{x}^i$ goes into the next generation if the following conditions are satisfied:

$$\|I - R\hat{x}^i\|_2 < \|I - Rx^i\|_2. \tag{7}$$

Evolution of the population corresponds to the dynamics of a "swarm of midges" (i. e. random point clouds). The cloud is moving along the relief of optimized function, repeating landscape features. In the case of falling, it takes the shape of the ravine and the points distribution is such that the expectation of the difference between two random vectors is directed along the long side of the ravine. This provides rapid movement along the narrow ravines. In similar conditions the gradient methods have vibrational dynamics "from wall to wall".

The pseudo-code for $\frac{1}{m}\|I - Rx\|_2$-minimization using differential evolution algorithm is shown below.

3 Empirical Results

3.1 Data description

In our empirical analysis we use publicly available data relating to three major market indices, that can be obtained from the OR-Library of [1, 2]. The three market indices are the Hang Seng (Hong Kong, $n = 31$), DAX 100 (Germany, $n = 85$) and the Nikkei 225 (Japan, $n = 225$) for $m = 290$ time periods each (weekly data), taken from [1]. The summary statistics of the daily log-returns of the indices are presented in Table 1. Table 1 shows that the return time series exhibit the typical patterns of financial times series: mean values around zero, light asymmetry and fat tails.

The data used in this paper is given in the form of matrices of asset prices. We transformed the original data sets into matrices of asset returns. It is widely accepted to use of the price ratio in order to derive the rate of returns, instead of using absolute asset price relations.

The index tracking problem with cardinality constraint were implemented using Matlab software, as well as built-in and specially developed functions. All simulations were run in Matlab. The system runs under MS Windows 10 64-bit and in our computational work we used an AMD FX-8350 pc with a 4.00 GHz processor and 8.0 GB RAM.

Algorithm 2: DIFFERENTIAL EVOLUTION ALGORITHM IN l_2

begin
- Generate N randomly distributed $y^i \in \mathbb{R}^n$, $i = 1, \ldots, N$;
- $\forall i$, set $y^i_j = 0$ for $n - K$ closest to 0 values of y^i_j;
- $\forall i$, set initial population P as $x^i_j = y^i_j / (\sum_{s=1}^n y^i_s)$, $j = 1, \ldots, n$;
- set L to the number of iterations;
- **while** $t \leq L$ **do**
  - **for** *each* x^i*,* $i = 1, \ldots, N$*, from* P **do**
    - select 3 random vectors x^a, x^b, x^c;
    - **for** *each* j *of* x^i_j **do**
      - with probability π_1: $z_{1,j} \leftarrow N(0, \sigma_1)$, else $z_{1,j} = 0$;
      - with probability π_2: $z_{2,j} \leftarrow N(0, \sigma_2)$, else $z_{2,j} = 0$;
      - $u_j \leftarrow U(0, 1)$;
      - **if** u_j ¡ $1 - \pi$ **then** $\tilde{x}^i_j = x^i_j$;
      - **else** $\tilde{x}^i_j = x^a_j + (F + z_{1,j})(x^b_j - x^c_j + z_{2,j})$;
      - $\forall \tilde{x}^i$, set $\tilde{x}^i_j = 0$ for $n - K$ closest to 0 elements;
      - $\forall \tilde{x}^i \in P$, set $\hat{x}^i = \tilde{x}^i / \sum_s \tilde{x}^i_s$;
      - **if** *conditions* (7) *are satisfied* **then** $\hat{x}^i$ replaces x^i in P;
- search $x^{i^*} = \arg\min_i \frac{1}{m} \|I - Rx^i\|_2$;
- **return** x^{i^*};

end

Table 1. Summary statistics of the indices' weekly returns

Data set	n	m	mean, %	std, %	skew	kurt	min	max
1 Hang Seng	31	290	0.42	3.32	-0.04	3.85	-0.12	0.11
2 DAX 100	85	290	0.25	2.03	-0.21	3.72	-0.07	0.07
5 Nikkei 225	225	290	-0.01	2.85	0.44	4.85	-0.11	0.12

3.2 Comparison using stochastic dominance approach

Description of stochastic dominance principle. To compare the approaches examined in this study we used stochastic dominance principle enabling a choice to be made in favour of one or another method (portfolio) [15]. Its peculiarity lies in the fact that it does not require precise knowledge of the investor's utility function, it only has to be monotonous and not decreasing [22].

Before proceeding to the definition of stochastic dominance, let us consider the concept of a dominant portfolio. The portfolio is considered to be dominant, i. e. preferred over other portfolio, if it has a higher level of return at the same level of risk or lower risk for the same expected return than another portfolio.

In this study, we will use stochastic dominance of first and second order for the analysis of the resulting portfolios. A more detailed description can be found in the book [19].

Let us denote by F_1 and F_2 distribution functions of random variables (portfolio returns) X_1 and X_2 accordingly. If on the interval $[a, b]$ (random variable support) the inequality $F_1 \leq F_2$ is satisfied, i. e. return distribution function of one portfolio does not exceed return distribution function of the other portfolio, we can say that there is stochastic dominance of the first order of one portfolio over the other (or random variable X_1 over X_2), and denote $X_1 \succeq_{\mathrm{I}} X_2$.

However, the situation may not be unambiguous, i. e. return distribution functions of the first and second portfolios may overlap. Therefore we can not say with certainty which of the portfolios is preferable for the investor. In this case the evaluation can be performed on the basis of stochastic dominance of the second order.

We say that this is the case of stochastic dominance of the second order of random variable X_1 over X_2 and use the notation $X_1 \succeq_{\mathrm{II}} X_2$, if

$$\int_a^x F_1(y)dy \leq \int_a^x F_2(y)dy, x \in [a, b].$$

Thus, the criterion of stochastic dominance of the second order is based not on the comparison of portfolio return distribution functions but on the integrals of these functions, i. e. areas under distribution functions. It can also be said that the first portfolio is preferable to the second if the cumulative distribution function of its return never exceeds, and at least in one case is less than cumulative distribution function of the second portfolio.

As follows from the second-order dominance, dominance of the first order of one portfolio over the other automatically assumes its stochastic dominance of the second order as well. Thus, the condition of the second-order stochastic dominance is a weaker condition.

In summary, we can note that in comparison with other methods of assessment stochastic dominance gives the investor a more general approach to the assessment of risky portfolios.

We determine the optimal model using a window of 100 observations (weeks) and leave it intact for further 10 out-of-sample trading weeks for testing purposes. Then, this (in-sample) window is shifted forward by 10 weeks, and a new portfolio of solution for index tracking problem is determined using a window of the new 100 observations, and then again it is left unchanged for further 10 out-of-sample weeks, and so on. Thus, portfolios are recalculated once every 10 weeks. It should be noted that the comparison of models for in-sample data was held on the last 10 observations (in-sample) of the 100-observations window. It was done so to ensure that in-sample and out-of-sample samples were of the same dimension.

To compare these approaches in terms of stochastic dominance, all in-sample and out-of-sample data were joined. As a result, two time series of returns had been received with respect to the index of 190 weeks in length, respectively for in-sample and out-of-sample data. Stochastic dominance principle was used for such time series.

Comparative analysis of greedy algorithm and LASSO-type algorithm. Table 2 shows comparative results of stochastic dominance for greedy algorithm and LASSO-type algorithm. It should be noted that stochastic dominance of the first order was not observed for all samples. While stochastic dominance (of the second order) was observed only for 4 out of 6 out-of-sample data (S&P 100 and Nikkei 225 for both indicators δ) in favour of greedy algorithm.

Table 2. Comparison results of LASSO-type (abbrev. L) and the greedy algorithm (abbrev. G) for three data sets

	$\delta = 0.9$		$\delta = 0.25$	
	In-Sample	Out-of-Sample	In-Sample	Out-of-Sample
Hang Seng	5 to 10 stocks		12 to 17 stocks	
	—	—	—	—
S&P 100	8 to 16 stocks		19 to 34 stocks	
	—	$G \succeq_{II} L$	—	$G \succeq_{II} L$
Nikkei 225	10 to 20 stocks		31 to 40 stocks	
	—	$G \succeq_{II} L$	—	$G \succeq_{II} L$

For clarity, Fig. 1(a) demonstrates the comparison of return distribution functions for portfolios built on greedy algorithm and LASSO-type algorithm for out-of-sample data Nikkei 225 ($\delta = 0.25$). The graph shows that the probability distribution functions overlap, i. e. stochastic dominance of the first order is not possible. Fig. 1(b) represents the comparison of *cumulative* distribution functions for the same data. The graph shows that accumulated distribution function of greedy algorithm is always lower than accumulated distribution function of LASSO-type algorithm, i. e. we have stochastic dominance of the second order, $X_{\text{Greedy}} \succeq_{\text{II}} X_{\text{LASSO}}$.

The following two figures 1(c) and 1(d) show an example for out-of-sample data Hang Seng ($\delta = 0.25$), when we can not claim the presence of stochastic dominance. Namely, both probability distribution functions and cumulative return distribution functions of the portfolios built using greedy algorithm and LASSO-type algorithm overlap.

Comparative analysis of greedy algorithm and differential evolution algorithm. Table 3 represents comparative results of stochastic dominance for greedy algorithm and differential evolution algorithm. It can be seen that stochastic dominance of the first order was not observed in all samples. If we compare the algorithms in terms of stochastic dominance of the second order, we will see that preference is given to differential evolution algorithm (in 1 out of 6 in-sample data, and in 3 out of 6 out-of-sample data the portfolio built on differential evolution algorithm stochastically dominates over (the second order) portfolio built on greedy algorithm). While inverse stochastic dominance is observed only in 1 case for in-sample data, and in 1 case for out-of-sample data

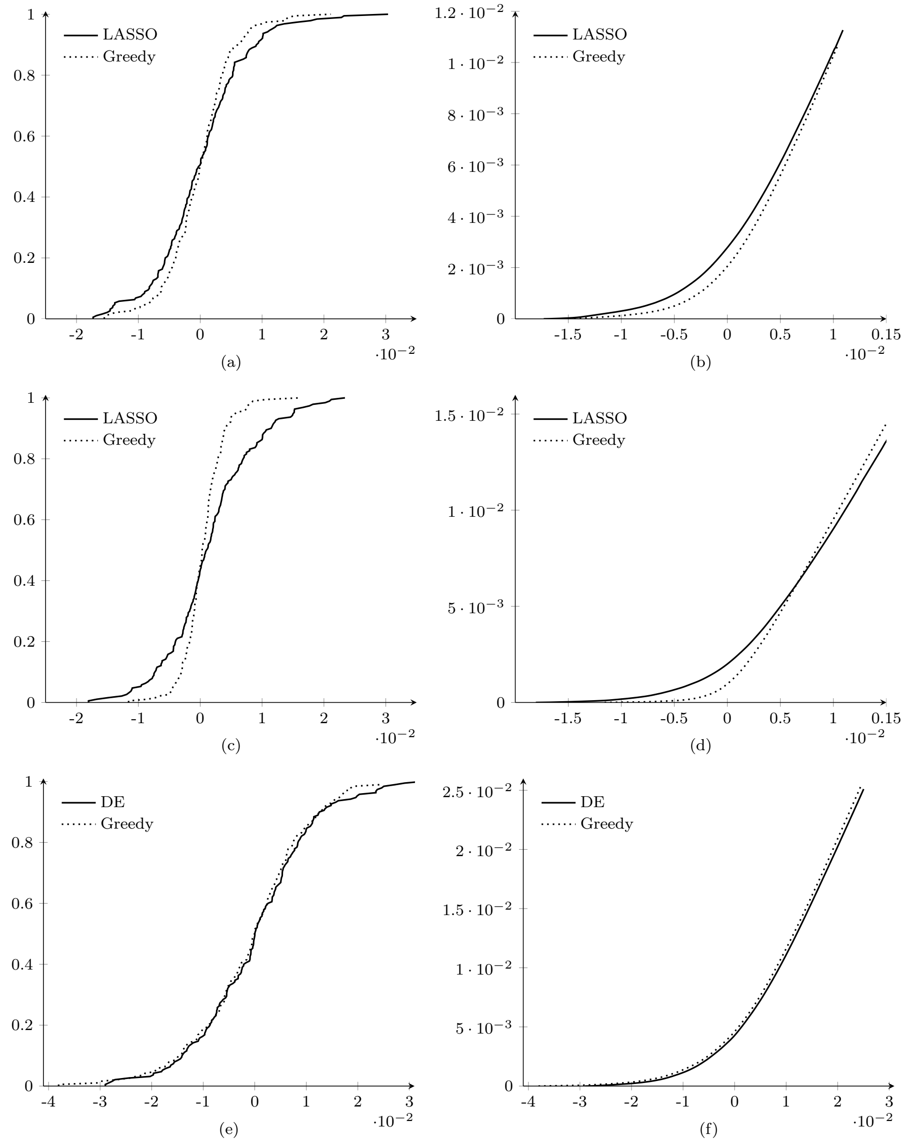

Fig. 1. Return distribution functions and cumulative return distribution functions

(Nikkei 225 for $K = 20$). It is possible that differential evolution algorithm in this case was worse due to insufficient number of operations, or populations, which would allow it to produce a more accurate solution.

By way of illustration, Fig. 1(e) demonstrates the comparison of cumulative return distribution functions of the portfolios built on greedy algorithm and differential evolution algorithm for out-of-sample data Nikkei 225 ($K = 5$). The

Table 3. Comparison results of the differential evolution (abbrev. DE) and the greedy algorithm (abbrev. G) for three data sets

	$K = 5$		$K = 20$	
	In-Sample	Out-of-Sample	In-Sample	Out-of-Sample
Hang Seng	$DE \succeq_{\mathrm{II}} G$	$DE \succeq_{\mathrm{II}} G$	—	—
S&P 100	—	—	—	$DE \succeq_{\mathrm{II}} G$
Nikkei 225	—	$DE \succeq_{\mathrm{II}} G$	$G \succeq_{\mathrm{II}} DE$	$G \succeq_{\mathrm{II}} DE$

graph ahows that probability distribution functions overlap, i. e. there is no first-order stochastic dominance. The following figure 1(f) shows comparison of accumulated distribution functions for the same data. The graph demonstrates that accumulated distribution function of greedy algorithm is always higher than accumulated distribution function of differential evolution algorithm, i. e. we have stochastic dominance of the second order ($X_{\mathrm{DE}} \succeq_{\mathrm{II}} X_{\mathrm{Greedy}}$).

4 Conclusion

Summing up the results of comparison of greedy algorithm and LASSO-type algorithm, we should note that in most cases we can not give preference to one or the other portfolio. However, for 4 out of 12 data sets yet there was the second-order stochastic dominance in favour of greedy algorithm. Also it is important to note that greedy algorithm is significantly easier to implement, its time costs are small and it easily copes with cardinality as compared to the LASSO-type algorithm.

As for greedy algorithm and differential evolution algorithm comparison, it should be said that the portfolios built on greedy algorithm and differential evolution algorithm are not significantly different in the selection of assets, and, as a rule, without using short sales. Moreover, though the portfolios built on differential evolution algorithm as a whole stochastically dominate over (the second-order) portfolios built on greedy algorithm, greedy algorithms significantly surpass differential evolution algorithms in terms of ease of implementation and algorithm execution time.

References

1. Beasley, J., Meade, N., Chang, T.J.: An evolutionary heuristic for the index tracking problem. European Journal of Operational Research 148(3), 621–643 (2003)
2. Beasley, J.E.: Or-library: distributing test problems by electronic mail. Journal of the Operational Research Society pp. 1069–1072 (1990)
3. Becker, S., Candès, E.J., Grant, M.C.: Templates for convex cone problems with applications to sparse signal recovery. Math. Program. Comput. 3(3), 165–218 (2011)

4. Brodie, J., Daubechies, I., De Mol, C., Giannone, D., Loris, I.: Sparse and stable Markowitz portfolios. Proceedings of the National Academy of Sciences of the USA 106(30), 12267–12272 (2009)
5. Canagkoz, N., Beasley, J.: Mixed-integer programming approaches for index tracking and enhanced indexation. European Journal of Operational Research 196(1), 384–399 (2008)
6. Chang, T.J., Meade, N., Beasley, J., Sharaiha, Y.: Heuristics for cardinality constrained portfolio optimisation. Computers and Operations Research 27(13), 1271–1302 (2000)
7. Das, A., Kempe, D.: Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In: Getoor, L., Scheffer, T. (eds.) Proceedings of the 28th International Conference on Machine Learning (ICML-11). pp. 1057–1064. ACM, New York, NY, USA (June 2011)
8. Derigs, U., Nickel, N.H.: Meta-heuristic based decision support for portfolio optimization with a case study on tracking error minimization in passive portfolio management. OR Spectrum 25, 345–378 (2003)
9. Gilli, M., Kellezi, E.: Financial Engineering, E-Commerce, and Supply Chain, chap. The threshold accepting heuristic for index tracking, pp. 1–18. Kluwer Academic, Dordrecht (2002)
10. Gilli, M., Winker, P.: Handbook of Computational Econometircs, chap. Heuristic optimization methods in econometrics, pp. 81–120. Wiley: Chichester (2009)
11. Jeurissen, R., van den Berg, J.: Index tracking using a hybrid genetic algorithm. In: Computational Intelligence Methods and Applications, 2005 ICSC Congress on. pp. 1–6 (2005)
12. Jeurissen, R., van den Berg, J.: Optimized index tracking using a hybrid genetic algorithm. In: Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on. pp. 2327–2334 (June 2008)
13. Le Thi, H.A., Moeini, M.: Long-short portfolio optimization under cardinality constraints by difference of convex functions algorithm. J. Optim. Theory Appl. 161(1), 199–224 (Apr 2014)
14. Li, Y., Yang, X., Zhu, S.S., Li, D.H.: A hybrid approach for index tracking with practical constraints. Journal of Industrial and Management Optimization 10(3), 905–927 (2014)
15. Linton, O., Maasoumi, E., Whang, Y.: Consistent testing for stochastic dominance under general sampling schemes. Review of Economic Studies 72(3), 735–765 (2005)
16. Maringer, D., Oyewumi, O.: Index tracking with constrained portfolios. Intell. Syst. Account., Finance Mgmt 15, 57–71 (2007)
17. Maringer, D.: Portfolio Management with Heuristic Optimization (Advances in Computational Management Science). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
18. Peng, Z.: The optimisation on the multi-period mean-average absolute deviation portfolio selection in friction market. International Journal of Intercultural Information Management 2(4), 343–352 (2011)
19. Prigent, J.L.: Portfolio Optimization and Performance Analysis. Chapman & Hall/CRC, Boca Raton (2007)
20. Roll, R.: A Mean/Variance Analysis of Tracking Error. The Journal of Portfolio Management 18(4), 13–22 (1992)
21. Rudolf, M., Wolter, H.J., Zimmermann, H.: A linear model for tracking error minimization. Journal of Banking & Finance 23(1), 85–103 (1999)
22. Scaillet, O., Topaloglou, N.: Testing for stochastic dominance efficiency. Journal of Business & Economic Statistics 28(1), 169–180 (2010)

23. Sidorov, S., Faizliev, A., Khomchenko, A.: Algorithms for l1-norm minimisation of index tracking error and their performance. Int. J. Mathematics in Operational Research X(Y), XX–XX (2016)
24. Storn, R., Price, K.: Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11(4), 341–359 (1997)
25. Takeda, A., Niranjan, M., ya Gotoh, J., Kawahara, Y.: Simultaneous pursuit of out-of-sample performance and sparsity in index tracking portfolios. Computational Management Science 10(1), 21–49 (2013)
26. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society (Series B) 58, 267–288 (1996)
27. Woodside-Oriakhi, M., Lucas, C., Beasley, J.E.: Heuristic algorithms for the cardinality constrained efficient frontier. European Journal of Operational Research 213(3), 538–550 (2011)
28. Zhang, P., Zhang, W.G.: Multiperiod mean absolute deviation fuzzy portfolio selection model with risk control and cardinality constraints. Fuzzy Sets and Systems 255, 74–91 (2014)