Computational Statistics and Data Analysis 76 (2014) 737–759

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Solving norm constrained portfolio optimization via coordinate-wise descent algorithms

Yu-Min Yen [a,*], Tso-Jung Yen [b]

[a] Institute of Economics, Academa Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan

[b] Institute of Statistical Science, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 11529, Taiwan

ARTICLE INFO

Article history:
Received 5 October 2012
Received in revised form 3 July 2013
Accepted 3 July 2013
Available online 9 August 2013

Keywords:
Minimum variance portfolio
Weighted norm constraint
Berhu penalty
Grouped portfolio selection

ABSTRACT

A fast method based on coordinate-wise descent algorithms is developed to solve portfolio optimization problems in which asset weights are constrained by l_q norms for $1 \leq q \leq 2$. The method is first applied to solve a minimum variance portfolio (mvp) optimization problem in which asset weights are constrained by a weighted l_1 norm and a squared l_2 norm. Performances of the weighted norm penalized mvp are examined with two benchmark data sets. When the sample size is not large in comparison with the number of assets, the weighted norm penalized mvp tends to have a lower out-of-sample portfolio variance, lower turnover rate, fewer numbers of active constituents and shortsale positions, but higher Sharpe ratio than the one without such penalty. Several extensions of the proposed method are illustrated; in particular, an efficient algorithm for solving a portfolio optimization problem in which assets are allowed to be chosen grouply is derived.

1. Introduction

How to select assets to form an optimal portfolio is one of the central issues in financial studies. To solve a portfolio selection problem such as the mean–variance portfolio optimization (Markowitz, 1952), we usually need to estimate mean vector and covariance matrix of the asset returns, and then plug the estimates into the optimization problem. If there are N assets, then the total number of parameters needed to be estimated for the mean vector and covariance matrix is $N+N(N+1)/2$. Accurate estimations for these parameters are necessary for successfully implementing a portfolio selection strategy; however, it is not an easy task, especially when N becomes large. If sample size n for estimating these parameters is not relatively large enough to N, cumulative estimation errors of these estimated parameters will become non-negligible, and the optimal mean–variance portfolio with these calibrations will fail to work. Empirical evidences on bad performances of the mean–variance portfolio strategy due to insufficient sample size can be found in Jagannathan and Ma (2003), DeMiguel et al. (2009a) and Kan and Zhou (2007). Kan and Smith (2008) showed that when the ratio N/n is not small enough, if simple sample estimations are used, they generally will cause an upward biasness on mean and downward biasness on variance of return of the optimal portfolio. Consequently the resulting in-sample estimation on Sharpe ratio will be too optimistic.

To reduce impacts from the estimation errors, we can choose a smaller number of assets, say $N' < N$, and at the same time optimize the objective function in the portfolio optimization. Selecting fewer assets for a portfolio means that the optimal portfolio weight vector should have some elements exactly equal to zero. Such a portfolio is termed sparse portfolio in Brodie et al. (2009). Ideally, the sparse portfolio may be obtained by solving the following l_0 norm constrained minimum variance

* Corresponding author.
E-mail addresses: yyu_min@hotmail.com, YMYEN@econ.sinica.edu.tw (Y.-M. Yen), tjyen@stat.sinica.edu.tw (T.-J. Yen).

0167-9473/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.csda.2013.07.010

portfolio (mvp) optimization problem:

$$\min_{\mathbf{w}} \mathbf{w}^{\mathbf{T}} \Sigma \mathbf{w} \text{ subject to } \|\mathbf{w}\|_{l_0} \leq N', \mathbf{w}^{\mathbf{T}} \mathbf{1}_N = 1, \text{ possibly } \mathbf{w}^{\mathbf{T}} \boldsymbol{\mu} = \overline{\mu}, \tag{1}$$

where $\|\mathbf{w}\|_{l_0} = \sum_{i=1}^{N} \mathbb{I}\{w_i \neq 0\}$, and $\mathbb{I}\{A\}$ denotes the indicator function such that $\mathbb{I}\{A\} = 1$ if event A is true and $\mathbb{I}\{A\} = 0$ otherwise. The $N \times N$ matrix Σ is the covariance matrix of the N asset returns. Throughout the paper, we assume Σ is symmetric and positive semidefinite (psd). In practice, Σ can be any kind of psd covariance matrix estimation. The $N \times 1$ vector $\boldsymbol{\mu} = (\mu_{1,\ldots,}\mu_N)^{\mathbf{T}}$ is the vector of expected asset returns, and $\overline{\mu}$ is the investor's required return.

Solving (1) involves combinatorial optimization, and it becomes extremely difficult when N is large. Practically we can replace the l_0 norm $\|\mathbf{w}\|_{l_0}$ with the l_1 norm $\|\mathbf{w}\|_{l_1} := \Sigma_{i=1}^{N}|w_i|$ in (1). The l_1 norm constraint can facilitate zero components (sparsity) in the weight vector, and hence it can function as the l_0 norm constraint to restrict the number of assets in the portfolio. The l_1 norm constraint also is a convex function of $\mathbf{w}$, and such convex relaxation makes the modified portfolio optimization problem easily tractable even when N becomes very large.

In statistics, the l_1 norm penalty approach is frequently used in dealing with high dimensional estimation problems. For example, Tibshirani (1996) proposed the lasso, which aims to regularize OLS estimation with the l_1 norm penalty on regression coefficients. The l_1 norm penalty also has been widely used on estimating structures of networks (Meinshausen and Buhlmann, 2006; Vinciotti and Hashem, 2013). For portfolio optimization problems, imposing the l_1 norm constraint can improve portfolio performances when the number of assets becomes very large (Brodie et al., 2009; DeMiguel et al., 2009a; Fan et al., 2012; Welsch and Zhou, 2007). In addition, the portfolio optimization may be affected by changes of the estimated parameters, and imposing the l_1 norm constraint can help to mitigate the effects and stabilize the optimization (Brodie et al., 2009; Fan et al., 2012).

In this paper, we develop fast and easy-to-implement coordinate-wise descent algorithms to solve the norm constrained portfolio optimization problems. We first focus on an algorithm for solving the optimal minimum variance portfolio (mvp) with a weighted l_1 and squared l_2 norm penalty and linear constraints, and later we will show the proposed method can be extended to mvp optimization problems with various norm constraints. The algorithms previously used to solve such norm constrained portfolio optimization problems are either quadratic programming or the least angle regression (LARS) type algorithms (Efron et al., 2004). Recently, the coordinate-wise descent algorithms have been shown to be powerful tools for solving large dimensional variable selection problems in which the norm penalties are imposed on covariate coefficients (Friedman et al., 2007). We demonstrate that the coordinate-wise descent algorithms can also be used to solve various norm constrained portfolio optimization problems.

The norm constraints also have been adopted on the index tracking problem in which a portfolio is formulated to replicate a market index. For example, Giamouridis and Paterlini (2010) used the l_1 norm constraint on the problem and showed that it results in better out-of-sample tracking performances. Gotoh and Takeda (2011) discussed the relations between norm constraints and robust portfolio optimization problems and then applied an approach of norm constrained conditional value-at-risk (CVaR) portfolio optimization to the index tracking problem. Fastrich et al. (in press) proposed to use the nonconvex l_q norm penalty, $0 < q < 1$ on the index tracking problem. In Takeda et al. (2013), they proposed to use both the l_0 and squared l_2 norm penalties on tracking a stock index. The authors developed a greedy algorithm to solve the penalized portfolio optimization and then applied their method on tracking Nikkei 225 index.

In addition to the norm constraint approach, we can assign asset weights with some simple rules in order to avoid massive estimations. The value weighted and equally weighted ($1/N$) portfolios are such examples. DeMiguel et al. (2009b) showed how such simple strategies can outperform more sophisticated strategies. Another frequently used way is to construct more robust statistical estimators for the mean vector and covariance matrix of the asset returns, such as bias-adjusted or Bayesian shrinkage estimators (El Karoui, 2009; Jorion, 1986; Kan and Zhou, 2007; Ledoit and Wolf, 2003; Lai et al., 2011), and use them in the portfolio optimization problems. We also can combine the improved portfolios to form a new portfolio; for example Frahm and Christoph (2010) and Tu and Zhou (2011) showed that a suitable linear combination of weights of a benchmark portfolio and a more sophisticated strategy often provides better performances than either only one of them is considered. It is natural to incorporate the latter two approaches with the norm constraint strategy.

The rest of the paper is organized as follows. In Section 2, we introduce a benchmark case of the mvp optimization in which the asset weights are constrained by the weighted l_1 and squared l_2 norm. We then describe a coordinate-wise descent algorithm for solving the benchmark case in Section 3. In Section 4, we use real data sets to examine empirical properties of the weighted norm mvp. In Section 5, we discuss some extensions, which include portfolio optimization problems with different convex norm penalties, and possible ways to use our method in portfolio optimization when nonconvex norm constraints are imposed (e.g., Fastrich et al. (2012b)) or even more general objective functions are considered. We also have a brief discussion on limitations of our method. Section 6 is conclusion.

2. Weighted norm minimum variance portfolio

To begin our analysis, we first consider the minimum variance portfolio (mvp) optimization in which the asset weights are constrained by the weighted l_1 and squared l_2 norm constraint:

$$\min_{\mathbf{w}} \mathbf{w}^{\mathbf{T}} \Sigma \mathbf{w} \text{ subject to } \alpha \, \|\mathbf{w}\|_{l_1} + (1-\alpha) \, \|\mathbf{w}\|_{l_2}^2 \leq c \text{ and } \mathbf{w}^{\mathbf{T}} \mathbf{1}_N = 1,$$