ÖREBRO UNIVERSITY
SCHOOL OF BUSINESS

WORKING PAPER

3/2019

An Iterative Approach to Ill-Conditioned Optimal Portfolio Selection

Mårten Gulliksson and Stepan Mazur
Statistics

ISSN 1403-0586

http://www.oru.se/Institutioner/Handelshogskolan-vid-Orebro-universitet/Forskning/Publikationer/Working-papers/
Örebro University School of Business
701 82 Örebro
SWEDEN

An Iterative Approach to Ill-Conditioned Optimal Portfolio Selection

MÅRTEN GULLIKSSON[a] AND STEPAN MAZUR[b,1]

[a] *School of Science and Technology, Örebro University, SE-70182 Örebro, Sweden*

[b] *School of Busniess, Örebro University, SE-70182 Örebro, Sweden*

Abstract

Covariance matrix of the asset returns plays an important role in the portfolio selection. A number of papers is focused on the case when the covariance matrix is positive definite. In this paper, we consider portfolio selection with a singular covariance matrix. We describe an iterative method based on a second order damped dynamical systems that solves the linear rank-deficient problem approximately. Since the solution is not unique, we suggest one numerical solution that can be chosen from the iterates that balances the size of portfolio and the risk. The numerical study confirms that the method has good convergence properties and gives a solution as good as or better than the constrained least norm Moore-Penrose solution. Finally, we complement our result with an empirical study where we analyze a portfolio with actual returns listed in S&P 500 index.

JEL Classification: C10, C44

Keywords: Mean-variance portfolio, singular covariance matrix, linear ill-posed problems, second order damped dynamical systems

[1]Corresponding author. E-mail address: Stepan.Mazur@oru.se. Stepan Mazur acknowledges financial support from the internal research grants at Örebro University, and from the project "Models for macro and financial economics after the financial crisis" (Dnr: P18-0201) funded by Jan Wallander and Tom Hedelius Foundation.

1 Introduction

Modern portfolio theory has drawn much attention in the academic literature starting from 1952 when Harry Max Markowitz published his seminal paper about portfolio selection (see Markowitz (1952)). He proposed efficient way of portfolio allocation that guarantees the lowest risk for a given level of the expected return.

A number of papers are devoted to questions like, e.g., how can an optimal portfolio be constructed, monitored, and/or estimated by using historical data (see, e.g., Alexander & Baptista (2004), Golosnoy & Okhrin (2009), Bodnar (2009), Bodnar et al. (2017a), Bauder et al. (2018)), what is the influence of parameter uncertainty on the portfolio performance (cf., Okhrin & Schmid (2006), Bodnar & Schmid (2008)), how do the asset returns influence the portfolio choice (see, e.g., Jondeau & Rockinger (2006), Mencia & Sentana (2009), Adcock (2010), Harvey et al. (2010), Amenguala & Sentana (2010)), how is it possible to estimate the characteristics of the distribution of the asset returns (see, e.g., Jorion (1986), Wang (2005), Frahm & Memmel (2010)), how can the structure of optimal portfolio be statistically justified (Gibbons et al. (1989), Britten-Jones (1999), Bodnar & Schmid (2009)). Björk et al. (2014) studied the mean-variance portfolio optimization in continuous time, whereas Liesiö & Salo (2012) developed a portfolio selection framework which uses the set inclusion to capture incomplete information about scenario probabilities and utility functions. Chiarawongse et al. (2012) formulated a mean-variance portfolio selection problem that accommodates qualitative input about expected returns and provided an algorithm that solves the problem, while Levy & Levy (2014) analyzed the parameter estimation error in portfolio optimization.

All above discussed papers are focused on the case when the covariance matrix of the asset returns is positive definite. In practice, covariance matrix is unknown, therefore, it needs to be estimated using historical data. Most common estimators are sample and maximum likelihood estimators. If the number of observations is greater than the number of assets in the portfolio then both estimators of the covariance matrix are positive definite. However, if the number of observations is less than the number of assets in the portfolio then both estimators of the covariance matrix are singular. Since optimal portfolio weights depend on the inverse of the covariance matrix, the original optimization problem formulated by Markowitz (1952) will have an infinite number of solutions. In Pappas et al. (2010), the solution to the optimization problem with singular covariance matrix is obtained by replacing the inverse with the Moore-Penrose inverse. It leads us to the unique solution with the minimal Euclidean norm. Statistical properties of the optimal portfolio weights for small sample and singular covariance matrix are well studied by Bodnar et al. (2016, 2017b) and Bodnar et al. (2018).

The main aim of the present paper is to deliver an alternative approach that works better than Moore-Penrose inverse. In particular, we employ an iterative method that solves the linear ill-posed problem approximately. Iterative methods for linear ill-posed problems are certainly not new and include classical methods like Landweber iteration with Nesterov acceleration and Conjugate gradient methods, see Neubauer (2000, 2017). Very recently a new approach has been developed based on second order damped dynamical systems, see Zhang & Hofmann (2018) for the ill-posed linear case. In Zhang & Hofmann (2018) it is shown that the method is a regularization method, i.e., loosely speaking for an ill-posed linear problem there exists a unique solution when the number of iterations tend to infinity and the error tends to zero. We have applied and extended this method to the rank-deficient and linearly constrained portfolio selection problem considered here. Specifically, we show that the method is convergent and how to choose optimal parameters (time step and damping). As seen in Sections 3.1, 3.2 the iterative method generally performs better in the sense of giving a smaller risk and smaller weights of the optimal portfolio.

The rest of the paper is structured as follows. In Section 2, an iterative approach to ill-conditioned optimal portfolio selection is discussed. The results of numerical and empirical studies are discussed in details in Section 3, while Section 4 summarizes the paper.

2 Main Results

We consider a portfolio of k assets and let $\mathbf{x}_i = (x_{1i}, \ldots, x_{ki})^T$ be the k-dimensional vector of log-returns of these assets at time $i = 1, \ldots, N$. We assume that the second moment of $\mathbf{x}_i$ is finite. Let the mean vector of the asset returns be denoted by $\boldsymbol{\mu}$ and the covariance matrix by $\boldsymbol{\Sigma}$ such that $rank(\boldsymbol{\Sigma}) = r \leq k$, i.e. $\boldsymbol{\Sigma}$ can also be singular matrix. Let $\mathbf{w} = (w_1, \ldots, w_k)^T$ be the vector of portfolio weights, where w_j denotes the weight of the jth asset, and let $\mathbf{1}$ be the vector of ones while $\mathbf{I}$ be the identity matrix. In general we allow for short sales and therefore for negative weights.

The classical problem of portfolio selection is defined as

$$\min_{\mathbf{w}} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^T \mathbf{1} = 1, \ \mathbf{w}^T \boldsymbol{\mu} = q \tag{1}$$

where q is the expected rate of return that is required on the portfolio. If $\boldsymbol{\Sigma}$ is a positive definite matrix then the optimization problem (1) has unique solution which is given by

$$\mathbf{w} = \frac{C - qB}{AC - B^2} \boldsymbol{\Sigma}^{-1} \mathbf{1} + \frac{qA - B}{AC - B^2} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \tag{2}$$

where $A = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}$, $B = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, $C = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$. However, if $\boldsymbol{\Sigma}$ is singular then the optimization problem (1) has an infinite number of solutions. Pappas et al. (2010) suggested the solution that appears to be unique with the minimal Euclidean norm and is obtained by replacing the inverse with the Moore-Penrose inverse

$$\mathbf{w} = \frac{C - qB}{AC - B^2}\boldsymbol{\Sigma}^+\mathbf{1} + \frac{qA - B}{AC - B^2}\boldsymbol{\Sigma}^+\boldsymbol{\mu} \tag{3}$$

with $A = \mathbf{1}^T\boldsymbol{\Sigma}^+\mathbf{1}$, $B = \mathbf{1}^T\boldsymbol{\Sigma}^+\boldsymbol{\mu}$, $C = \boldsymbol{\mu}^T\boldsymbol{\Sigma}^+\boldsymbol{\mu}$.

In practice, both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown parameters and the investor cannot determine $\mathbf{w}$. Consequently, she/he should estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ using previous observations. The most common estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are given by

$$\hat{\boldsymbol{\mu}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i = \frac{1}{N}\mathbf{X}\mathbf{1}, \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x}_i - \hat{\boldsymbol{\mu}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}})^T = \frac{1}{N-1}\mathbf{X}\mathbf{V}\mathbf{X}^T, \tag{4}$$

where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, and $\mathbf{V} = \mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T$ is a symmetric and idempotent matrix, i.e., $\mathbf{V} = \mathbf{V}^T$ and $\mathbf{V}^2 = \mathbf{V}$. If $N \geq k$ then sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ is positive definite, but $\hat{\boldsymbol{\Sigma}}$ is singular when $N < k$. Hence, we can get the case when portfolio size is larger the sample size and, therefore, sample covariance matrix will be singular. Then one can use the solution with smallest Euclidean norm that is defined in (3). Alternatively, one can use an iterative approach that is discussed in the next section.

Here we will derive a new unconstrained problem by projecting $\mathbf{w}$ to the subspace of the constraints. Define

$$\mathbf{B} = \begin{bmatrix} \mathbf{1}^T \\ \boldsymbol{\mu}^T \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 1 \\ q \end{bmatrix},$$

then the constraints can be written as $\mathbf{B}\mathbf{w} = \mathbf{c}$ and the solution is $\mathbf{w} = \mathbf{Z}\mathbf{u} + \mathbf{g}$ where $\mathbf{Z}$ spans the null space of $\mathbf{B}$ and $\mathbf{g}$ is any solution to $\mathbf{B}\mathbf{w} = \mathbf{c}$. After some algebra the minimization problem (1) can be written as

$$\min_{\mathbf{u}} \mathbf{u}^T\mathbf{Z}^T\boldsymbol{\Sigma}\mathbf{Z}\mathbf{u} + 2\mathbf{g}^T\boldsymbol{\Sigma}\mathbf{Z}\mathbf{u} + \mathbf{g}^T\boldsymbol{\Sigma}\mathbf{g} = \Phi(\mathbf{u}). \tag{5}$$

The solution will be uniquely defined if $\mathbf{M} = \mathbf{Z}^T\boldsymbol{\Sigma}\mathbf{Z}$ is invertible. In the following we always choose $\mathbf{g} = \mathbf{B}^+\mathbf{c}$ where $\mathbf{B}^+$ is the Moore-Penrose inverse of $\mathbf{B}$, i.e., $\mathbf{g}$ is the least norm solution to the constraint equations.

2.1 The Discrete Functional Particle Method

Let us first consider the minimization problem

$$\min_{\mathbf{u}\in\mathcal{R}^n} V(\mathbf{u}), \tag{6}$$

where $V : \mathcal{R}^n \to \mathcal{R}$ is at least a twice continuously differentiable convex function giving a unique solution $\mathbf{u}^*$. We use the conventional notation for inner-product in $\mathcal{R}^n$ and norm as $(\cdot, \cdot)$ and $\|\cdot\|$, respectively, with subindices added to specify the underlying norm if needed.

The main idea for solving (6) is to utilize the fact that the solution $\mathbf{u}^*$ to (6) is also a stationary solution to the second order damped dynamical system

$$\ddot{\mathbf{u}}(t) + \eta\dot{\mathbf{u}}(t) = -\nabla V(\mathbf{u}(t)),\ \eta > 0, \tag{7}$$

and this solution is unique and globally exponentially stable, see, e.g., references in Bégout et al. (2015). The problem (7) naturally appears in modelling mechanical systems where a relevant example is the heavy ball with friction system (HBF). In this case (7) describes the motion of a material point with positive mass (equal to 1 in (7)). The optimization properties for HBF with different friction have been studied in detail in Attouch et al. (2000); Attouch & Alvarez (2000); Alvarez (2000), and references therein. Throughout the whole chapter we reserve the dot notation for the derivatives with respect to the fictious time t.

After the problem has been formulated as (7), it is necessary to choose a numerical method for solving (7) that is efficient and fast enough. The key observation to make this choice is to recognize that the dynamical system (7) is Hamiltonian with the total energy

$$E(u(t), \dot{\mathbf{u}}(t)) = V(\mathbf{u}(t)) + \frac{1}{2}\|\dot{\mathbf{u}}(t)\|^2, \tag{8}$$

which, in fact, is also a Lyapunov function to (7) (and, as stated before, therefore guarantees convergence). It is easy to see that the energy will decrease exponentially in time resulting in an exponential decay of $\|\mathbf{u}(t) - \mathbf{u}^*\|$, Bégout et al. (2015). Symplectic methods, such as symplectic Runge-Kutta, Störmer-Verlet, and etc., are tailor-made for Hamiltonian systems and preserve the energy (Bhatt et al. (2016)). This serves as the motivation for our choice of numerical method.

Let us rewrite (7) as the first order system

$$\begin{aligned} \dot{\mathbf{u}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= -\eta\mathbf{v} - \nabla V(\mathbf{u}). \end{aligned} \tag{9}$$

Then, we apply a one step symplectic explicit method, such as symplectic Euler or Störmer-Verlet Hairer et al. (2006), which give us an iterative map on the form

$$\mathbf{w}_{k+1} = F(\mathbf{w}_k, \Delta t_k, \eta_k), \mathbf{w}_k = (\mathbf{u}_k, \mathbf{v}_k), k = 1, 2, \ldots, \tag{10}$$

where the time step Δt_k, damping η_k may be independent of k. The choice of parameters Δt_k and η_k can be aimed to optimize the performance of the numerical method, which generally is a non-trivial task.

We call the approach of finding the solution to (6) by solving (7) with a symplectic method as the *dynamical functional particle method* (DFPM) (see Gulliksson et al. (2013)).

We would like to emphasize that it is the combination of the damped dynamical system together with an efficient (fast, stable, accurate) symplectic solver that makes DFPM a very competitive method.

Let us now return to our main optimization problem (1) and assume that we have eliminated the constraints as in (5). If we define $\mathbf{M} = \mathbf{Z}^T\mathbf{\Sigma}\mathbf{Z}, \mathbf{d} = \mathbf{Z}^T\mathbf{\Sigma}\mathbf{g}$ we get the unconstrained problem

$$\min_{\mathbf{u}\in\mathcal{R}^{n\times n-2}} \frac{1}{2}\mathbf{u}^T\mathbf{M}\mathbf{u} - \mathbf{u}^T\mathbf{d}, \quad \mathbf{M}\in\mathcal{R}^{n\times n-2}, \mathbf{d}\in\mathcal{R}^n, \tag{11}$$

where we will assume that $\mathbf{M}$ is positive semidefinite. In the general setting of the minimization problem (6) we can define

$$V(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T\mathbf{M}\mathbf{u} - \mathbf{u}^T\mathbf{d} \tag{12}$$

and it is straightforward to formulate DFPM for $V(\mathbf{u})$ in (12) as

$$\ddot{\mathbf{u}} + \eta\dot{\mathbf{u}} = \mathbf{d} - \mathbf{M}\mathbf{u}. \tag{13}$$

or the first order system

$$\begin{aligned} \dot{\mathbf{u}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= -\eta\mathbf{v} + (\mathbf{d} - \mathbf{M}\mathbf{u}). \end{aligned} \tag{14}$$

Additionally we need intial conditions, say, $\mathbf{u}(0) = \mathbf{u}_0$, $\dot{\mathbf{u}}(0) = \mathbf{v}_0$. Using symplectic Euler on (14) we get

$$\begin{cases} \mathbf{v}_{k+1} = (I - \Delta t\,\eta)\mathbf{v}_k - \Delta t(\mathbf{d} - \mathbf{M}\mathbf{u}_k) \\ \mathbf{u}_{k+1} = \mathbf{u}_k + \Delta t\,\mathbf{v}_{k+1} \end{cases} \tag{15}$$

or, equivalently,

$$\mathbf{w}_{k+1} = \mathbf{G}\mathbf{w}_k + c,\ \mathbf{G} = \begin{bmatrix} \mathbf{I} - \Delta t^2\mathbf{M} & \Delta t(1-\Delta t\,\eta)\,\mathbf{I} \\ \Delta t\,\mathbf{I} & (1-\Delta t\,\eta)\,\mathbf{I} \end{bmatrix},\ \mathbf{g} = \begin{bmatrix} -\Delta t^2\mathbf{d} \\ -\Delta t\mathbf{d} \end{bmatrix}, \tag{16}$$

where $\mathbf{w}_k = [\mathbf{u}_k,\, \mathbf{v}_k]^T$.

It has been shown, see, e.g., Edvardsson et al. (2015) that DFPM based on (15) clearly outperforms classical iterative methods such as Gauss-Seidel, Jacobi, as well as the method based on a first order ODE. Conjugate gradient methods can in some cases compete in efficiency with DFPM but do not have the regularization properties that we need here, see Zhang & Hofmann (2018) for details.

2.2 Convergence analysis and choice of parameters

When $\mathbf{M}$ in (11) has full rank $\mathbf{u}(t)$ will converge to the unique solution $\mathbf{u} = \mathbf{M}^{-1}\mathbf{d}$. Turning to the the rank-deficient case we consider the SVD of $\mathbf{M} = \mathbf{U}\boldsymbol{\Sigma}_M\mathbf{U}^T$ and tranform (13) to

$$\ddot{\mathbf{y}} + \eta\dot{\mathbf{y}} = \mathbf{f} - \boldsymbol{\Sigma}_M\mathbf{y}, \boldsymbol{\Sigma}_M = \mathrm{diag}(s_1, \ldots, s_{r_M}, 0, \ldots, 0), s_i > 0, \tag{17}$$

where r_M is the rank of $\mathbf{M}$. Since the system now is decoupled we can by partitioning $\mathbf{f} = \left[\mathbf{f}_1^T, \mathbf{f}_2^T\right]^T$ write the solution of (17) as

$$\begin{aligned} y_i(t) &= \frac{1}{s_i}f_i^1 + \alpha_i^{11}e^{-\gamma_i^1 t} + \alpha_i^{12}e^{-\gamma_i^2 t},\ i = 1, \ldots, r_M, \\ y_i(t) &= \alpha_i^{21} + \alpha_i^{12}e^{-\eta * t} + \frac{1}{\eta}f_i^2 t,\ i = r_M + 1, \ldots, n, \end{aligned}$$

where $\gamma_i^j = \eta/2 \pm \sqrt{\eta^2/4 - s_i}$ and α_i^j are given by the initial conditions. We conclude that the solution is unbounded and grows linearly with t. However, that does not mean that the dynamical system (17) can not be used for getting one, of infinitely many, solutions to (1). Indeed, by iteratively solving (17) and carefully choosing when to stop the iterations we attain a, hopefully useful, regularized solution. This is called iterative regularization in the literature of ill-posed problems and is an alternative to the least norm solution Vogel (2002).

In order to ensure fast convergence of the iterative scheme (15) one must choose the time step and damping such that for convergence $\|\mathbf{G}\| < 1$ and for efficiency $\|\mathbf{G}\|$ is as small as possible. The otpimal choice of parameters can be summarized in the following theorem, see Gulliksson (2017) and references therein.

Theorem 1. *Consider symplectic Euler (15) where $\lambda_i = \lambda_i(\mathbf{M}) > 0$ are the eigenvalues of $\mathbf{M}$. Then the parameters*

$$\Delta t = \frac{2}{\sqrt{\lambda_{min}} + \sqrt{\lambda_{max}}}, \eta = \frac{2\sqrt{\lambda_{min}}\sqrt{\lambda_{max}}}{\sqrt{\lambda_{min}} + \sqrt{\lambda_{max}}}, \tag{18}$$

where $\lambda_{min} = \min_i \lambda_i$ and $\lambda_{max} = \max_i \lambda_i$ are the solution to the problem

$$\min_{\Delta t, \eta} \max_{1 \le i \le 2n} |\mu_i(\mathbf{G})|,$$

where $\mu_i(\mathbf{G})$ are the eigenvalues of G.

However, the result in Theorem 1 is not applicable for the case when $\mathbf{M}$ does not have full rank since the smallest eigenvalue will be zero and the damping will in turn be zero giving an oscillating solution that does not converge to a solution of the underdetermined system. Therefore, we choose the parameters only in the subspace defined by the nonzero eigenvalues of $\mathbf{M}$.

Theorem 2. *Assume that* $\lambda_i(\mathbf{M}) > 0, i = 1, \ldots, r_M$ *are the nonzero eigenvalues of* $\mathbf{M}$ *and the rest of the eigenvalues are zero. Then symplectic Euler (15) with parameters*

$$\Delta t = \frac{2}{\sqrt{\lambda_{r_M}} + \sqrt{\lambda_{max}}}, \eta = \frac{2\sqrt{\lambda_{r_M}}\sqrt{\lambda_{max}}}{\sqrt{\lambda_{r_M}} + \sqrt{\lambda_{max}}}, \tag{19}$$

where $\lambda_{r_M} = \min_{\lambda_i > 0} \lambda_i$ *and* $\lambda_{max} = \max_i \lambda_i$ *is convergent.*

Proof. The time step is smaller and the damping is larger than the parameters given by Theorem 1 and therefore the method is stable and convergent. □

3 Numerical and Empirical Studies

3.1 Numerical Study

In this section we examine the iterative approach which is proposed by us. In particular, we evaluate optimal portfolio weights, its norm and variance of portfolio return. All results are compared with the optimal portfolio weights which are obtained by using Moore-Penrose inverse.

In the simulation study, we take $q = 10$. Each element of $\boldsymbol{\mu}$ is uniformly distributed on [0,1]. We also take $\boldsymbol{\Sigma} = \mathbf{U}^T \boldsymbol{\Lambda} \mathbf{U}$, where $\boldsymbol{\Lambda}$ is the $k \times k$ diagonal matrix with the first r non-zero elements which are uniformly distributed on $(0, 1]$ and multiplied by 10, while the rest of the diagonal elements of $\boldsymbol{\Lambda}$ are taken to be 0, $\mathbf{U}$ is the $k \times k$ orthogonal matrix that is formed from the orthonormal eigenvectors of $\boldsymbol{\Lambda}$. The results are compared for several values of $k \in \{10, 50, 100, 150, 300\}$ and $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$.

The solution is not unique but one numerical solution can be chosen from the iterates that balances the size of the portfolio and the risk, see Figure 1. Therefore, the choice of convergence criteria is important in order to stop the iterations at a satisfactory solution. We have chosen to use a relative convergence criterion based on the projected objective function, i.e., we stop the iterations when

$$\epsilon_k < \text{tolerance}, \ \epsilon_k = \|\nabla\Phi(\mathbf{u}_k)\|/\Phi(\mathbf{u}_k)$$

where $\mathbf{u}_k$ is the approximation of the solution of (5) at iteration k and $\Phi(\mathbf{u})$ is defined in (5). The tolerance is set to 10^{-12} in order to get a small risk. Again refering to Figure 1, we note that a higher tolerance will give a higher risk but smaller size of portfolio (norm of $\mathbf{w}$). Maximum number of iterations is taken to be 10^4 but in the presented tables this is not achieved.

In Figures 2-6, we present optimal portfolio weights that are obtained by using DFPM approach and Moore-Penrose inverse. We observe that behaviour of the weights is quite different. In particular, for $k \in \{100, 150, 300\}$ we can observe that absolute values of

	$r = 0.1k$	$r = 0.4k$	$r = 0.6k$	$r = 0.9k$
$k = 10$	1.9054e+01 (1.2339e+01)	1.7855e+01 (8.9739e+01)	1.2845e+01 (1.3443e+01)	1.0683e+01 (1.1076e+01)
$k = 50$	2.1670e+01 (2.0385e+01)	1.7561e+01 (8.4677e+00)	1.7440e+01 (1.2035e+01)	1.3007e+01 (1.3529e+01)
$k = 100$	2.8944e+01 (1.3737e+01)	2.6371e+01 (1.0171e+01)	2.4013e+01 (7.2916e+00)	1.0808e+01 (5.5681e+00)
$k = 150$	3.3765e+01 (1.0464e+01)	2.6818e+01 (1.1532e+01)	2.3927e+01 (5.1976e+00)	1.5123e+01 (1.2805e+01)
$k = 300$	4.8978e+01 (4.3182e+01)	3.8724e+01 (1.1921e+01)	3.0789e+01 (5.2208e+00)	1.4582e+01 (5.2424e+00)

Table 1: Norm of the optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse (in parentheses) for $k \in \{10, 50, 100, 150, 300\}$ and $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$.

	$r = 0.1k$	$r = 0.4k$	$r = 0.6k$	$r = 0.9k$
$k = 10$	1.5248e+00 (1.9118e+02)	1.3861e+01 (2.8239e+04)	3.9043e+02 (7.7022e+02)	6.1116e+02 (6.4385e+02)
$k = 50$	1.9047e–10 (2.0215e+03)	2.6400e+01 (2.5573e+02)	2.4803e–03 (6.6129e+00)	2.8018e+01 (3.7443e+01)
$k = 100$	1.5568e–03 (7.5202e+02)	1.1609e–01 (9.1053e+01)	2.2761e+00 (2.7120e+01)	1.6548e+01 (4.1219e+01)
$k = 150$	2.6065e–02 (3.8750e+02)	9.9835e–04 (7.1027e+01)	5.7977e–03 (6.0913e+01)	3.5722e+00 (1.1957e+01)
$k = 300$	6.4618e–05 (3.4047e+01)	1.5064e–04 (1.5929e+01)	3.8282e–02 (1.3888e+01)	4.5099e+00 (7.1193e+00)

Table 2: Variance of portfolio return obtained by using DFPM approach and Moore-Penrose inverse (in parentheses) for $k \in \{10, 50, 100, 150, 300\}$ and $r \in \{0.1k, 0.4k, 0.6k, 0.9k\}$.

almost all weights obtained by DFMP approach are much larger than the ones obtained by using Moore-Penrose inverse.

In Table 1, we present the norm of the optimal portfolio weights. For $k = 10$ and $r \in \{4, 6, 9\}$, we can observe that the norm of the weights from DFMP approach is smaller than the one obtained from the Moore-Penrose inverse. For $k = 50$ our method gives smaller norm when $r = 45$ only, while Moore-Penrose inverse delivers smaller norm in all other cases.

In Table 2, we compare the variance of portfolio returns for both methods. In all considered cases, we can observe that DFMP approach shows better performance than Moore-Penrose inverse. So, we can conclude that DFMP approach delivers smaller the variance of portfolio returns, while the norm of the weights can be larger than the one obtained by using Moore-Penrose inverse.

3.2 Empirical Study

In this section the results of an empirical study are presented. It is shown how one can apply the theory from the previous sections to real data. We use weekly S&P 500 log returns of 440 stocks for the period from the 4th of May, 2007 to the 25th of January, 2013 resulting in 300 observations. Expected rate of return is taken to be 10, i.e. $q = 10$. In DFMP approach, the convergence criterion, the tolerance and the maximum number of iterations are taken as in Section 3.1.

First of all, we estimate mean vector and covariance matrix by using their empirical counterparts that are defined in (4). Since the number of stocks $k = 440$ is greater than the sample size $N = 300$, sample estimator of the covariance matrix $\mathbf{\Sigma}$ will be singular matrix with $rank(\mathbf{\Sigma}) = N - 1 = 299$, i.e. $r = 299$. In Figure 7, we present eigenvalues of the sample covariance matrix. We can observe that the first 299 eigenvalues and much larger than the rest of the eigenvalues. It confirms that the rank of the sample covariance matrix is 299.

Since we have estimated both mean vector and covariance matrix, we are able to construct optimal portfolio weights by using DFMP approach and Moore-Penrose inverse. In Figure 8, we deliver the plot of optimal portfolio weights obtained by both methods. We can observe quite different behaviour of the weights. Consequently, we can see that both methods suggest completely different investment strategies. We also get that the norm of optimal portfolio weights from the DFMP approach is 8.6291e+02, while the norm from the Moore-Penrose inverse is 8.9725e+02. Moreover, the variance of portfolio return obtained by using DFMP approach is 4.4371e–02 and it is much smaller than the variance obtained from the Moore-Penrose inverse that is 3.2392e+01. Consequently, DFMP approach delivers better objective function as well as the norm of the portfolio weights.

4 Summary

Modern portfolio theory plays an important role in economics and suggests an investment strategy that give the lowest risk for a given level of expected return. If covariance matrix of the asset returns is positive definite, then Markowitz' optimization problem has unique solution that depends on the mean vector and covariance matrix of the asset returns. In practice, both mean vector and covariance matrix need to be estimated. Hence, one can get an estimator of the covariance matrix which is singular and, therefore, optimization problem will not have a unique solution. In our paper, we delivered a new iterative approach (DFPM) that solves the constrained and rank-deficient portfolio selection problem approximately. The method is based on using symplectic solvers for a damped dynamical system that solves the optimization problem but the solution is generally different from the least norm Moore-Penrose solution. We showed how to determine the optimal time step and damping for symplectic Euler that give fast convergence using only matrix-vector multiplications in each iteration step. In the numerical study we examined DFMP and compared it with analytical solution that is based on the Moore-Penrose inverse. The results are compared for several values of the portfolio size and the rank of the covariance matrix. We observed that iterative and analytical approaches deliver quite different investment strategies. We also found that the norm of the weights in DFMP approach is smaller than the one obtained from Moore-Penrose inverse in some cases only, while the variance of portfolio return is always smaller in all considered cases. In the empirical study, we analyzed weekly S&P 500 log returns of 440 stocks with 300 observations. It is shown that DFMP approach guarantees smaller variance of portfolio return than analytical solution based on the Moore-Penrose inverse.

References

Adcock, C. J. (2010). Asset pricing and portfolio selection based on the multivariate extended skew-student-t distribution. *Annals of Operation Research*, 176(1), 221–234.

Alexander, G. J. & Baptista, M. A. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, 50, 1261–1273.

Alvarez, F. (2000). On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM Journal on Control and Optimization*, 38(4), 1102–1119.

Amenguala, D. & Sentana, E. (2010). A comparison of mean-variance efficiency tests. *Journal of Econometrics*, 154, 16–34.

Attouch, H. & Alvarez, F. (2000). The heavy ball with friction dynamical system for convex constrained minimization problems. *Lecture notes in economics and mathematical systems*, 481, 25–35.

Attouch, H., Goudou, X., & Redont, P. (2000). The heavy ball with friction method, i. the continuous dynamical system: Global exploration of the local minima of a real-valued function by asymptotic analysis of a dissipative dynamical system. *Communications in Contemporary Mathematics*, 02(01), 1–34.

Bauder, D., Bodnar, T., Mazur, S., & Okhrin, Y. (2018). Bayesian inference for the tangent portfolio. *International Journal of Theoretical and Applied Finance*, DOI: 10.1142/S0219024918500541.

Bégout, P., Bolte, J., & Jendoubi, M. A. (2015). On damped second-order gradient systems. *Journal of Differential Equations*, 259(7), 3115 – 3143.

Bhatt, A., Floyd, D., & Moore, B. E. (2016). Second order conformal symplectic schemes for damped Hamiltonian systems. *Journal of Scientific Computing*, 66(3), 1234–1259.

Björk, T., Murgoci, A., & Zhou, Y. (2014). Mean-variance portfolio optimization with state-dependent risk aversion. *Mathematical Finance*, 24(1), 1–24.

Bodnar, O. (2009). Sequential surveillance of the tangency portfolio weights. *International Journal of Theoretical and Applied Finance*, 12, 797–810.

Bodnar, T., Mazur, S., & Okhrin, Y. (2017a). Bayesian estimation of the global minimum variance portfolio. *European Journal of Operational Research*, 256(1), 292–307.

Bodnar, T., Mazur, S., & Podgórski, K. (2016). Singular inverse Wishart distribution and its application to portfolio theory. *Journal of Multivariate Analysis*, 143, 314–326.

Bodnar, T., Mazur, S., & Podgórski, K. (2017b). A test for the global minimum variance portfolio for small sample and singular covariance. *AStA Advances in Statistical Analysis*, 101, 253–265.

Bodnar, T., Mazur, S., Podgórski, K., & Tyrcha, J. (2018). *Tangency portfolio weights in small and large dimensions: estimation and test theory.* Technical report, Örebro University School of Business.

Bodnar, T. & Schmid, W. (2008). A test for the weights of the global minimum variance portfolio in an elliptical model. *Metrika*, 67, 179–201.

Bodnar, T. & Schmid, W. (2009). Econometrical analysis of the sample efficient frontier. *The European Journal of Finance*, 15, 317–327.

Britten-Jones, M. (1999). he sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54, 655–671.

Chiarawongse, A., Kiatsupaibul, S., Tirapat, S., & Van Roy, B. (2012). Portfolio selection with qualitative input. *Journal of Banking & Finance*, 36, 489–496.

Edvardsson, S., Neuman, M., Edström, P., & Olin, H. (2015). Solving equations through particle dynamics. *Computer Physics Communications*, 197, 169 – 181.

Frahm, G. & Memmel, C. (2010). Dominating estimators for minimum-variance portfolios. *Journal of Econometrics*, 159, 289–302.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57, 1121–1152.

Golosnoy, V. & Okhrin, Y. (2009). Flexible shrinkage in portfolio selection. *Journal of Economic Dynamics and Control*, 33, 317–328.

Gulliksson, M. (2017). The discrete dynamical functional particle method for solving constrained optimization problems. *Dolomites Research Notes on Approximation*, 10, 6–12.

Gulliksson, M., Edvardsson, S., & Lind, A. (2013). The Dynamical Functional Particle Method. *ArXiv e-prints.*

Hairer, E., Lubich, C., & Wanner, G. (2006). *Geometric Numerical Integration, 2nd ed.* Springer.

Harvey, C. R., Leichty, J. C., Leichty, M. W., & Muller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, 10, 469–485.

Jondeau, E. & Rockinger, M. (2006). Optimal portfolio allocation under higher moments. *European Financial Management*, 12, 29–55.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21, 293–305.

Levy, H. & Levy, M. (2014). The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research*, 234(2), 372–381.

Liesiö, J. & Salo, A. (2012). Scenario-based portfolio selection of investment projects with incomplete probability and utility information. *European Journal of Operational Research*, 217(1), 162–172.

Markowitz, H. (1952). Mean-variance analysis in portfolio choice and capital markets. *Journal of Finance*, 7, 77–91.

Mencia, J. & Sentana, E. (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *Journal of Econometrics*, 153, 105–121.

Neubauer, A. (2000). On Landweber iteration for nonlinear ill-posed problems in Hilbert scales. *Numerische Mathematik*, 85, 309–328.

Neubauer, A. (2017). On Nesterov acceleration for Landweber iteration of linear ill-posed problems. *Journal of Inverse and Ill-Posed Problems*, 25, 381–390.

Okhrin, Y. & Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics*, 134, 235–256.

Pappas, D., Kiriakopoulos, K., & Kaimakamis, G. (2010). Optimal portfolio selection with singular covariance matrix. *International Mathematical Forum*, 5(47), 2305–2318.

Vogel, C. R. (2002). *Computational Methods for Inverse Problems.* Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies*, 18, 673–705.

Zhang, Y. & Hofmann, B. (2018). On the second order asymptotical regularization of linear ill-posed inverse problems. *ArXiv e-prints.*

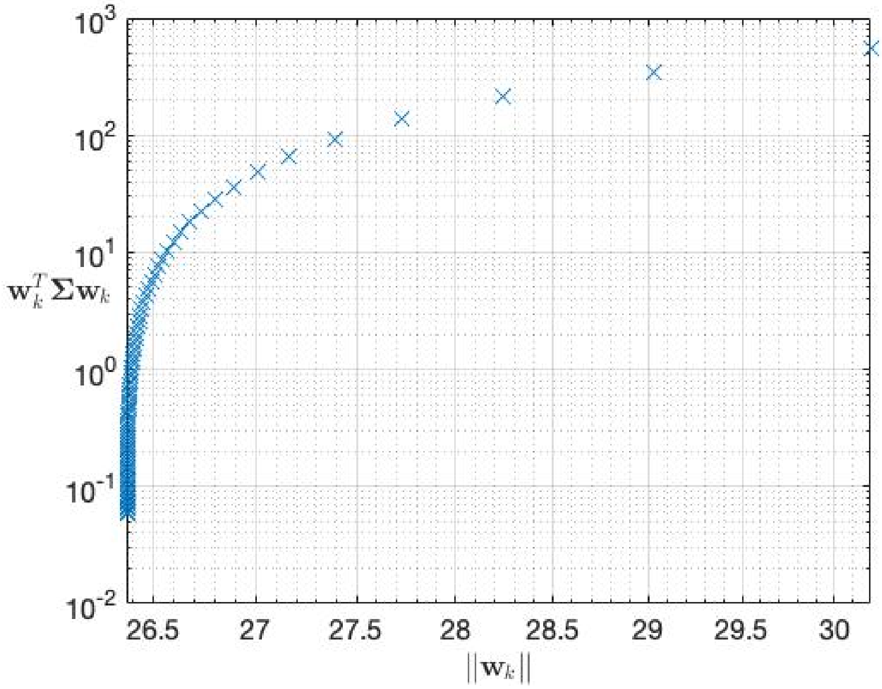
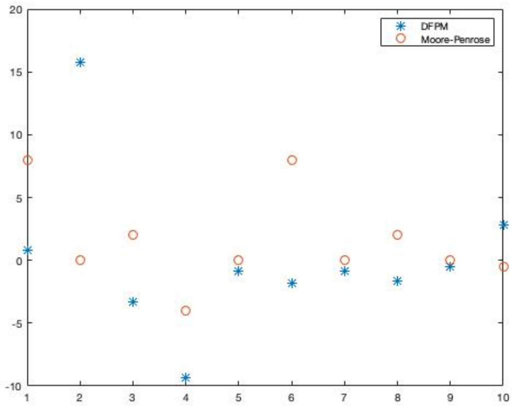
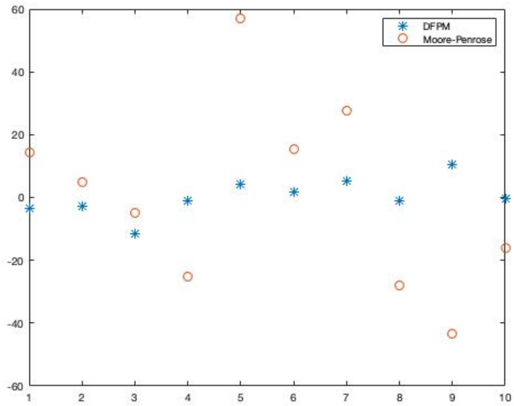


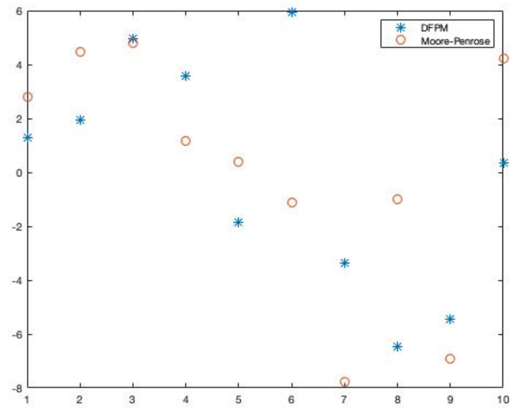
Figure 1: Plot of the risk as a function of size of portfolio weights (norm of $\mathbf{w}$) for $k = 100$ and $r = 40$.



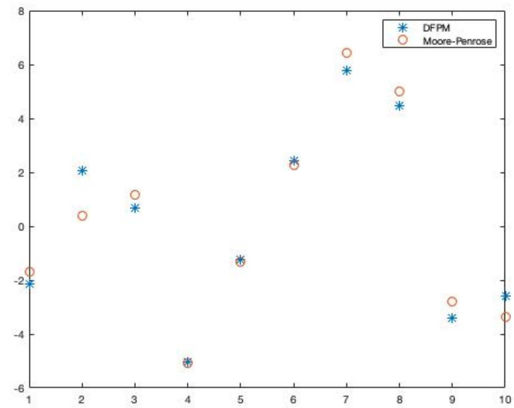
(a) $k = 10,\ r = 1$



(b) $k = 10,\ r = 4$

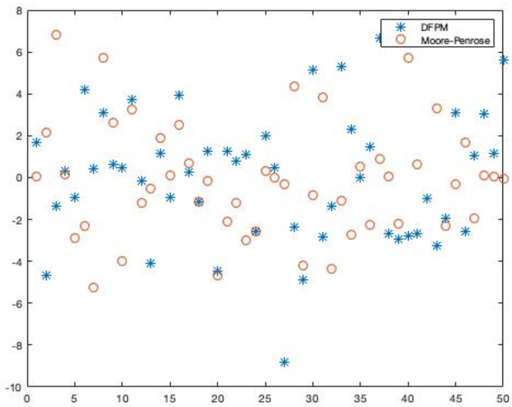


(c) $k = 10,\ r = 6$

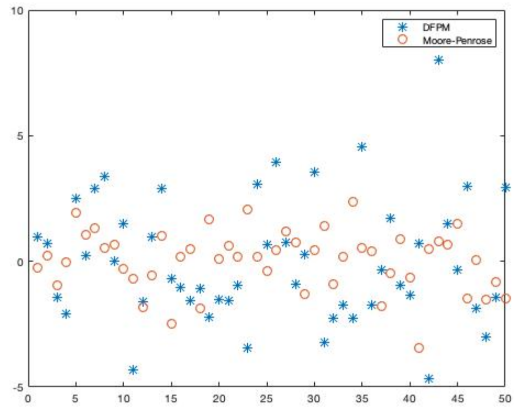


(d) $k = 10,\ r = 9$

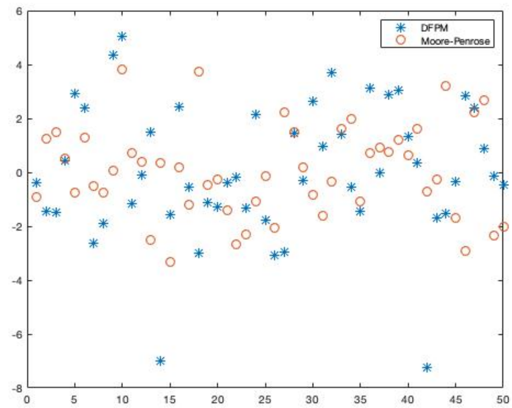
Figure 2: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 10$ and $r \in \{1, 4, 6, 9\}$.



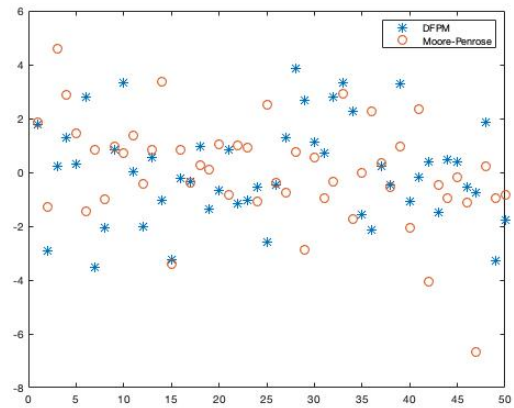
(a) $k = 50,\ r = 5$



(b) $k = 50,\ r = 20$

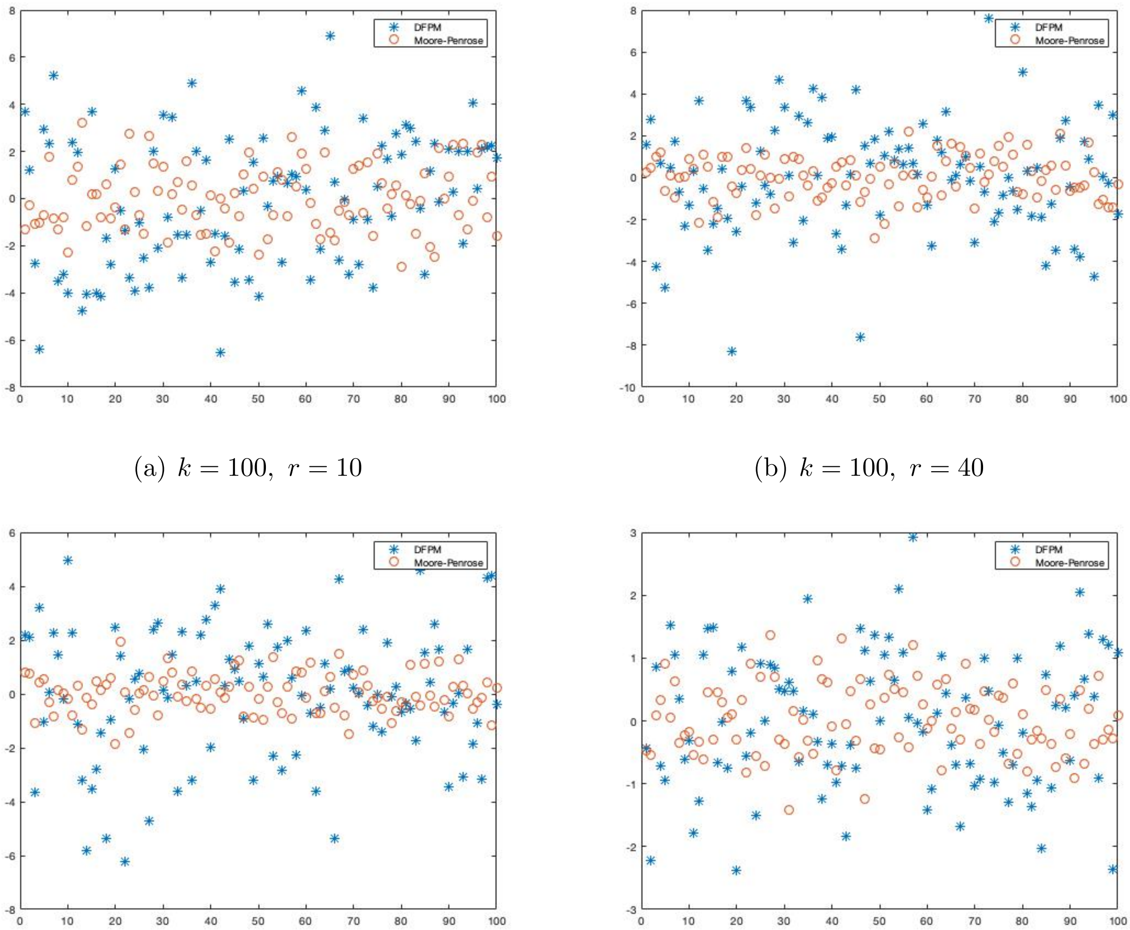


(c) $k = 50,\ r = 30$



(d) $k = 50,\ r = 45$

Figure 3: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 50$ and $r \in \{5, 20, 30, 45\}$.



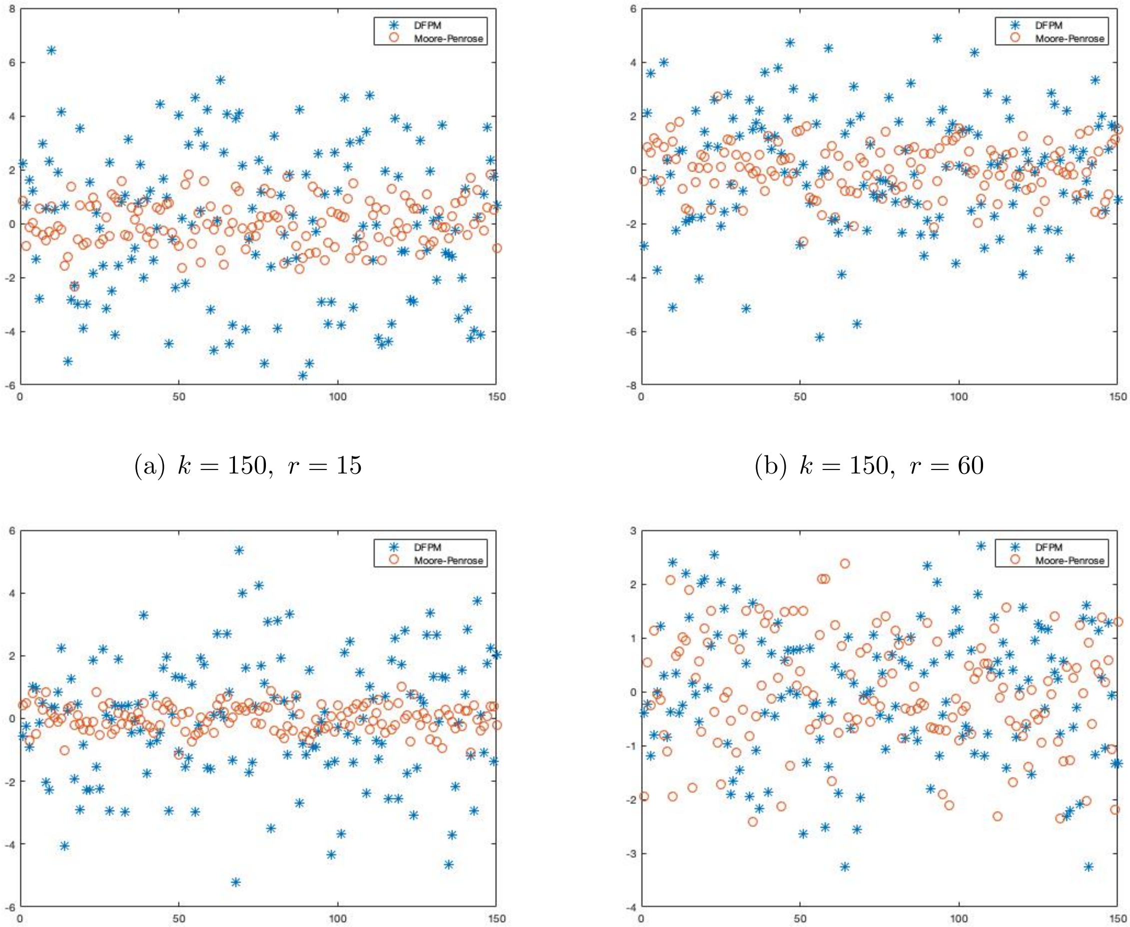
(a) $k = 100,\ r = 10$

(b) $k = 100,\ r = 40$

(c) $k = 100,\ r = 60$

(d) $k = 100,\ r = 90$

Figure 4: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 100$ and $r \in \{10, 40, 60, 90\}$.



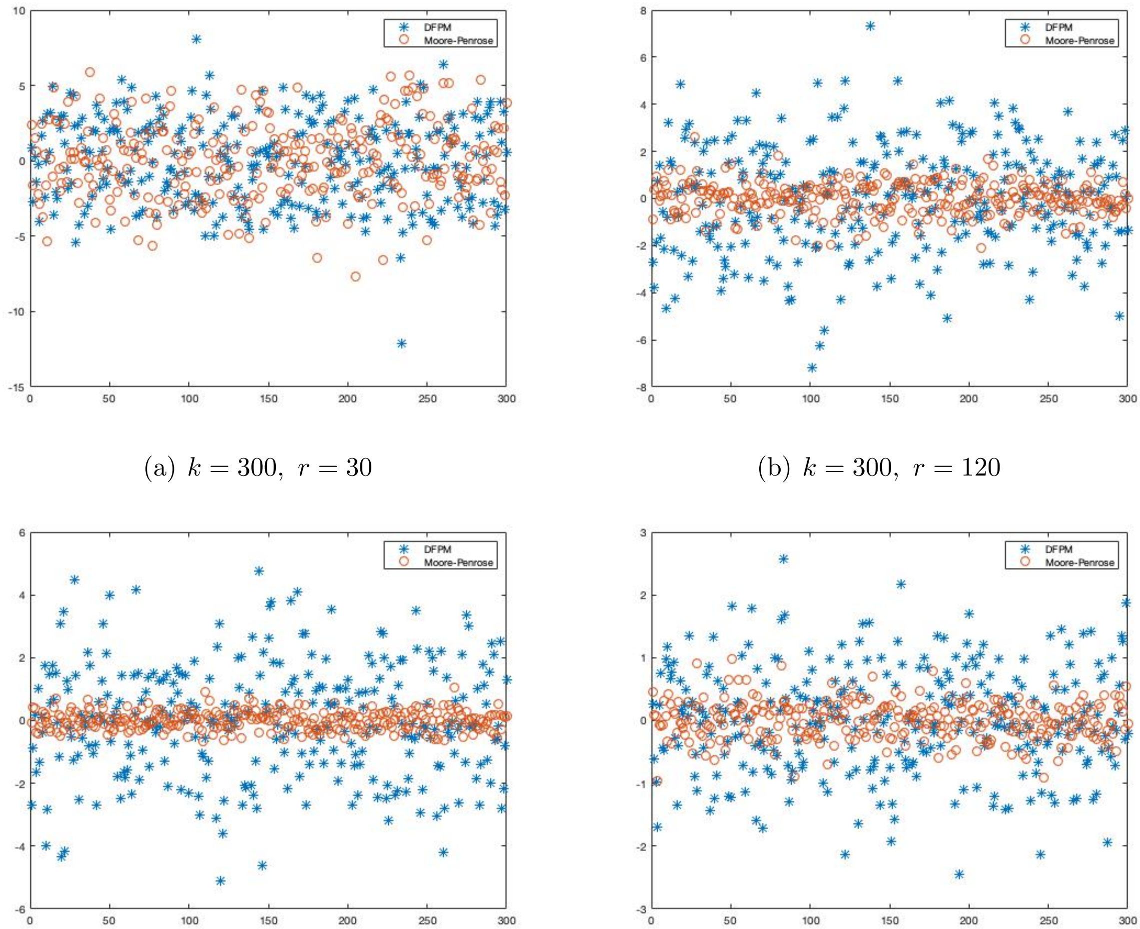
(a) $k = 150,\ r = 15$

(b) $k = 150,\ r = 60$

(c) $k = 150,\ r = 90$

(d) $k = 150,\ r = 135$

Figure 5: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 150$ and $r \in \{15, 60, 90, 135\}$.



(a) $k = 300,\ r = 30$

(b) $k = 300,\ r = 120$

(c) $k = 300,\ r = 180$

(d) $k = 300,\ r = 270$

Figure 6: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 300$ and $r \in \{30, 120, 180, 270\}$.

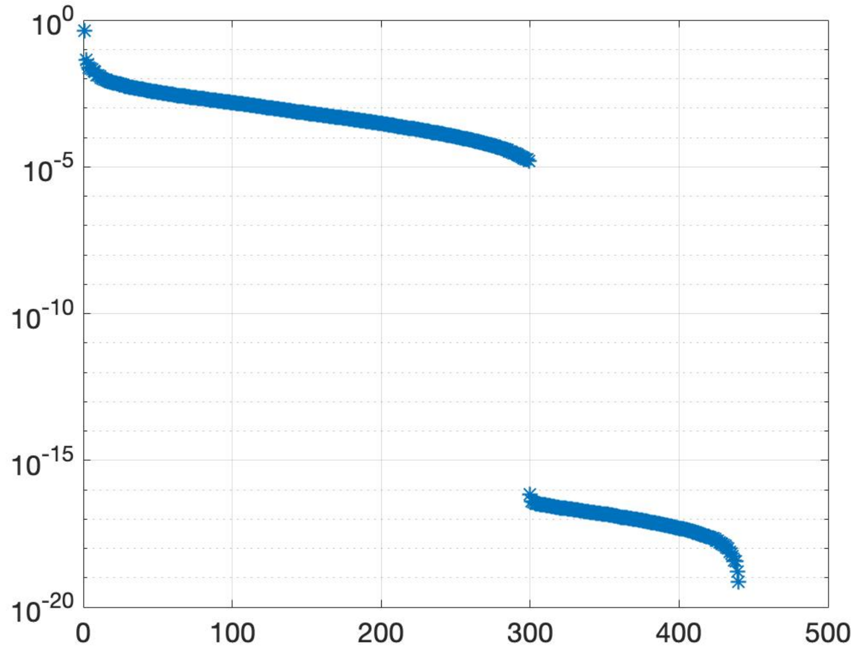


Figure 7: Eigenvalues of the sample covariance matrix for 440 stocks and 300 observations.

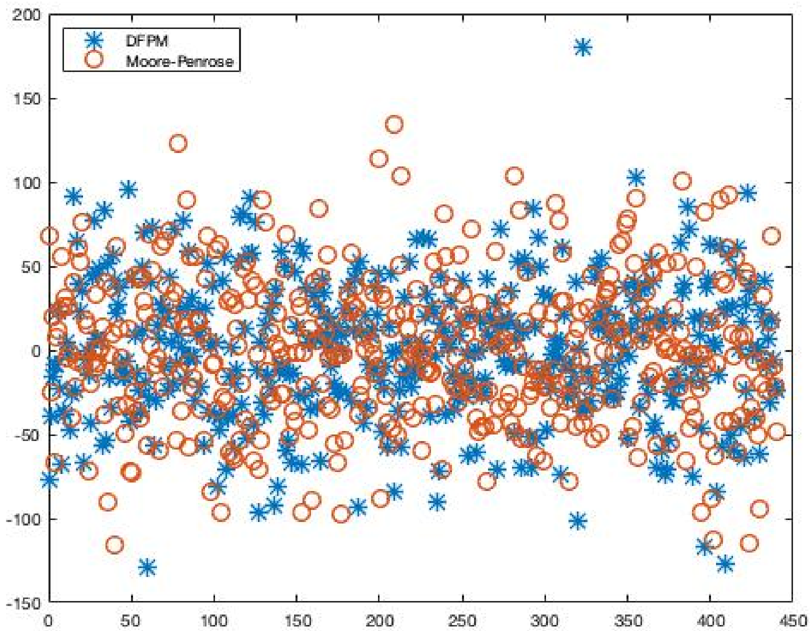


Figure 8: Optimal portfolio weights obtained by using DFPM approach and Moore-Penrose inverse for $k = 440$ and $r = 299$.