A New Portfolio Optimization Approach in the Singular Covariance Matrix: Improving Out-of-Sample Performance

Miyoung Lee [a], Jihun Kim [b], and Sekyung Oh [c*]

Abstract

This paper proposes a new portfolio optimization approach that does not rely on the covariance matrix and attains a higher out-of-sample Sharpe ratio than the existing approaches. Our approach is free from the problems related to the estimation of the covariance matrix, solves the corner solution problems of the Markowitz model in practice, improves the out-of-sample estimation of portfolio mean, and enhances the performance of portfolio by imposing certain structure on asset returns. Although the shrinkage to market estimator method shows the smallest out-of-sample standard deviation, it cannot perform the best in terms of Sharpe ratio when compared to our approach.

JEL Classifications: C13, G11, G15

Key words: Markowitz Model, Singular Problem, Shrinkage Estimator, Singular Value Decomposition, Out-of-Sample Sharpe Ratio

[a] Department of Business Administration, School of Business, ,Konkuk University, 120 Neungdong-ro, ,Gwangjin-gu, Seoul 05029, Korea. yura@konkuk.ac.kr

[b] Department of Business Administration, College of Business, KAIST, 85 Hoegi-ro, Dongdaemun-gu, Seoul 02455, South Korea; E-mail: jihunkim79@gmail.com

[c] Department of Business Administration, School of Business, ,Konkuk University, 120 Neungdong-ro, ,Gwangjin-gu, Seoul 05029, Korea. skoh@konkuk.ac.kr
*Corresponding author: Tel: 82-2-450-3641, Fax: 82-2-3436-6610

In his seminal paper, Markowitz [1952] proposed a portfolio optimization approach that minimizes the variance of a portfolio while achieving the expected return. To implement the Markowitz portfolio in practice, the means and covariances of asset returns need to be estimated. Traditionally, the sample means and covariances have been used but it is well known that the portfolios based on the sample estimates perform poorly out-of-sample due to estimation error (Frost and Savarino [1986, 1988] and Litterman [2003] among others). Since the estimation of means is more difficult and has a larger impact on out of sample performance of portfolio than the estimation of covariances, empirical financial economics with the applications of the Markowitz portfolio selection model have focused on minimum variance portfolio than mean variance portfolio (DeMiguel et al. [2008], Jagannathan and Ma [2003]).

The Markowitz model also poses other problems when the model is applied to actual data. When the number of observations T exceeds the number of assets N, the number of estimates to fill the covariance matrix grows exponentially as the number of assets increases or the errors in the estimation of parameters lead to unstable and extreme portfolio weights over time. Even worse, either when T is less than N, or when asset returns are linearly dependent, the covariance matrix becomes singular, in which case the Markowitz model does not work.

To solve the estimation error problem of the Markowitz model, several approaches have been developed. The heuristic approaches aim at producing more diversified portfolios either by introducing more constraints on portfolio weights (Frost and Savarino [1998], Jaganathan and Ma [2003]) or by averaging over portfolio weights that were obtained through the bootstrapping procedure (Jorion [1992], Michaud [1998]). On the other hand, the

Bayesian approaches directly adjust the inputs by combining prior information with sample data (Herold and Maurer [2006]).

There have been many attempts to find an invertible estimator of the covariance matrix. The pseudoinverse estimators of the covariance matrix are used by Sengupta [1983] and Pappas, Kiriakopoulos, and Kaimakamis [2010] or the shrinkage estimators of the covariance matrix are suggested by Ledoit and Wolf [2003]. Although both approaches find an invertible estimator of the covariance matrix of the Markowitz model one way or another, they have yet to solve the problem within the Markowitz framework that basically relies on the covariance matrix. In particular, Ledoit and Wolf [2003] propose to estimate the covariance matrix by an optimally weighted average of two existing estimators-the sample covariance matrix and single-index covariance matrix, and show their method achieves a significantly lower *out-of-sample variance* than a set of existing estimators. However, as they mention in the paper (p 605), they solely deal with the structure of risk in covariances, not with the structure of expected returns.

In this paper, we propose a new portfolio optimization approach that does not rely on the estimation of the covariance matrix and that attains a higher *out-of-sample Sharpe ratio* than the existing approaches. Since practitioners are generally looking for a portfolio with higher Sharpe ratios rather than that of lower variances, we need to use the out-of-sample Sharpe ratio rather than the out-of-sample variance to compare the performance of each portfolio optimization approach. The objective of the Markowitz model is to find the portfolio weights that minimize portfolio variance while achieving expected return *on average*. However, our approach attempts to find the portfolio that can achieve expected return *at each point in time* while minimizing the 2-norm of the portfolio weights.[1] By doing this, we can improve the estimation of means to result in higher out-of-sample Sharpe ratio of portfolio and solve the so-called corner solution problem when the Markowitz model is applied in practice.

Furthermore, the out-of-sample performance of our approach depends on how well the in-sample return vectors will span the out-of-sample return vectors. Since it is impossible to span the unsystematic risk part of the out-of-sample return vectors by the in-sample return vectors, we distinguish between the systematic risk part and the unsystematic risk part and attempt to span the out-of-sample systematic risk part by using the in-sample systematic part of return vectors alone. To impose some structure on the estimation of the systematic part of returns, we assume that stock returns are generated by Fama-French five factor model.

Our paper contributes to the literature on optimal portfolio choice in several ways. First, our approach is free from the problems related to the estimation of the covariance matrix because it does not need the covariance matrix when the portfolio weights are derived. Second, our approach solves the corner solution problems of the Markowitz model in practice because it minimizes the 2-norm of the portfolio weights and thus spreads out the weights. Third, our approach can improve the out-of-sample estimation of portfolio mean because it attempts to achieve target return at each point in time in-the-sample. Lastly, our approach can enhance the performance of portfolio in terms of out-of-sample Sharpe ratio and CEQ (Certainty-Equivalent) return by imposing certain structure on asset returns.

The remainder of this paper is organized as follows. The next section shows that the Markowitz approach to the portfolio optimization problem is basically the same as a system of linear equations for which the least squares method is applied. In the second section, we propose a new portfolio optimization approach that can contribute to solve the many problems that arise when the Markowitz model is applied in practice. In the third section, we explain data and show that new approach can greatly enhance the out-of-sample performance of portfolio. The last section concludes the paper.

MARKOWITZ MODEL REVISITED

Consider risky assets whose returns are r_{it} ($i = 1,2,\cdots,N$ and $t = 1,2,\cdots,T$). Denote the sample mean of each risky asset $\bar{r}_i$ for $i = 1,2,\cdots,N$ and the target return of the portfolio $\sum_{i=1}^{N} w_i \bar{r}_i = \text{q}$ (q: constant) with portfolio weights $w_i\text{'}s$, where $\sum_{i=1}^{N} w_i = 1$.The portfolio return at time t can be expressed as:

$$\sum_{i=1}^{N} w_i r_{it} = \text{q} + \varepsilon_t \text{, (t=}1,2,\cdots,T\text{)} \tag{1}$$

, where the $\varepsilon_t\text{'}$ s are residuals, $\sum_{i=1}^{N} w_i \bar{r}_i = \text{q}$, and $\sum_{i=1}^{N} w_i = 1$

Risk-averse investors would like to minimize the variance of Equation (1) while achieving their target return q. Then, the portfolio optimization problem can be written when short-selling is allowed as follows: [2]

$$\underset{w}{\text{Minimize}} \ \|\boldsymbol{\varepsilon}\| = \sqrt{\sum_{t=1}^{T} |\varepsilon_t|^2} \text{ ,} \tag{2}$$

$$\text{subject to } \ \boldsymbol{w}^T \boldsymbol{1} = 1, \textit{and} \ \ \boldsymbol{w}^T \bar{\boldsymbol{r}} = \text{q}$$

Here, $\|\boldsymbol{\varepsilon}\|$ means the Euclidean norm of a T-dimensional column vector $\boldsymbol{\varepsilon}$ of residuals, $\boldsymbol{w}^T$ represents the transpose of an N-dimensional column vector $\boldsymbol{w}$ of portfolio weights, and $\boldsymbol{1}$ and $\bar{\boldsymbol{r}}$ are the N-dimensional column vectors of the1's and $\bar{r}_i\text{'}s$, respectively.

We show that Equation (2) is equivalent to the portfolio optimization problem Markowitz tried to solve in his 1952 paper. If we substitute $\sum_{i=1}^{N} w_i \bar{r}_i$ for q in Equation (1) and rearrange the terms, then Equation (1) can be rewritten as:

$$\sum_{i=1}^{N} w_i \left(r_{it} - \bar{r}_i\right) = \varepsilon_t, (t = 1,2,\cdots,T) \tag{3}$$

Let $\boldsymbol{R}$ denote a $T \times N$ matrix whose ji^{th} element is $r_{ij} - \bar{r}_i$. Then, Equation (3) can be restated as follows using matrix notations:

$$\boldsymbol{Rw} = \boldsymbol{\varepsilon} \tag{4}$$

Since $\|\boldsymbol{\varepsilon}\|^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\boldsymbol{Rw})^T (\boldsymbol{Rw}) = \boldsymbol{w}^T (\boldsymbol{R}^T \boldsymbol{R}) \boldsymbol{w}$ and $\frac{1}{T-1} \boldsymbol{R}^T \boldsymbol{R} = \Sigma$, the portfolio optimization problem (2) can be rewritten as follows: [3]

$$\underset{w}{\text{Minimize}}\ \boldsymbol{w}^T \sum \boldsymbol{w} \tag{5}$$

$$\text{subject to } \boldsymbol{w}^T \boldsymbol{1} = 1, and\ \ \boldsymbol{w}^T \bar{\boldsymbol{r}} = \text{q}$$

, where $\boldsymbol{w}$ and $\boldsymbol{1}$ denote column vectors of weights of N assets and 1's, respectively; q is the target portfolio return; $\bar{\boldsymbol{r}}$ is a column vector of $\bar{r}_i$'s; and Σ is the covariance matrix of asset returns $\boldsymbol{r}_{it}$ $(i=1,2,\cdots, N, t=1,2,\cdots,T)$.

Equation (5) is exactly equivalent to the portfolio optimization problem formulated by Markowitz except for the non-negativity condition on the portfolio weights such that $w_i \geq 0, \forall i = 1,2,\cdots, N$, which is not necessary if short-selling is allowed. Also, because of Σ given in Equation (5), Markowitz assumed that there is an infinite number of time series data for the assets, which is impossible in reality (Markowitz [1952], Sengupta [1983]).

NEW APPROACH

Let $\boldsymbol{A}$ denote a $T \times N$ matrix whose ti^{th} element is r_{it} and $\boldsymbol{b}$ denote a T-dimensional vector defined as $\boldsymbol{b} = \text{q}\boldsymbol{1}$. Then, Equation (1) is rewritten as:

$$\boldsymbol{Aw} = \boldsymbol{b} + \boldsymbol{\varepsilon} \quad \text{or} \quad \boldsymbol{Aw} - \boldsymbol{b} = \boldsymbol{\varepsilon} \tag{6}$$

From Equation (6), we see that the Markowitz approach seeks to find $\overline{\boldsymbol{w}}$ that minimizes the error $\left\|\boldsymbol{Aw} - \boldsymbol{b}\right\|$ in the least squared sense.

EXHIBIT 1

Markowitz Problem and Least Squares Solution

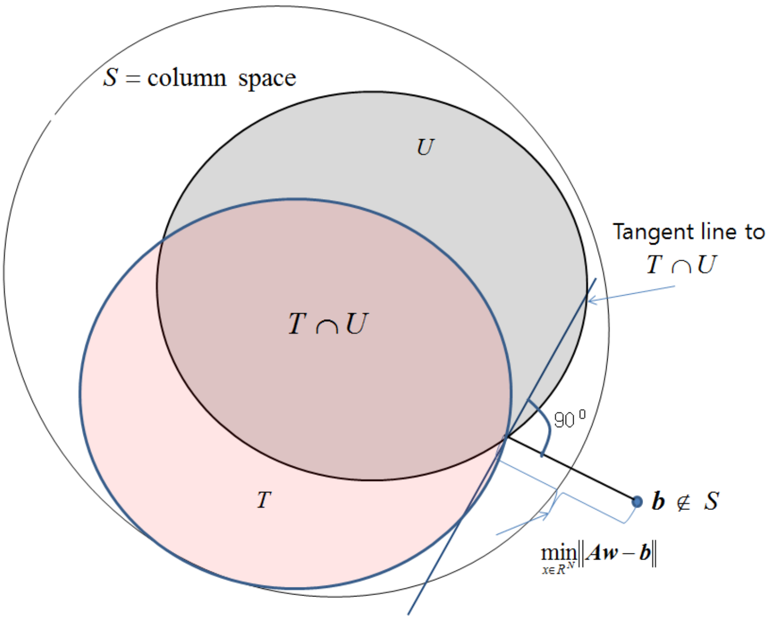

Exhibit 1shows that geometrically the error is exactly the distance from $\boldsymbol{b}$ to the point $\boldsymbol{Aw}$ in the column space S, where $S = Span\{\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_N\}$. Let's define U and T as follows.

$$\boldsymbol{U} = \{\boldsymbol{v} \in \boldsymbol{S} : \boldsymbol{v} = \sum_{i=1}^{N} \boldsymbol{w}_i \boldsymbol{A}_i \text{ with } \sum_{i=1}^{N} \boldsymbol{w}_i = 1\}, \quad \boldsymbol{T} = \left\{\boldsymbol{v} \in \boldsymbol{U} : \boldsymbol{v} = \sum_{i=1}^{N} \boldsymbol{w}_i \boldsymbol{A}_i \text{ with } \sum_{i=1}^{N} \boldsymbol{w}_i \bar{\boldsymbol{r}}_i = \text{q}\right\}$$

Then we can see that the tangent point which is in the set $\boldsymbol{T} \cap \boldsymbol{U}$ satisfying the constraints of the Markowitz model and is also orthogonal to $\boldsymbol{b}$ (which is not in S) is the solution to the Markowitz problem.

EXHIBIT 2

New Approach

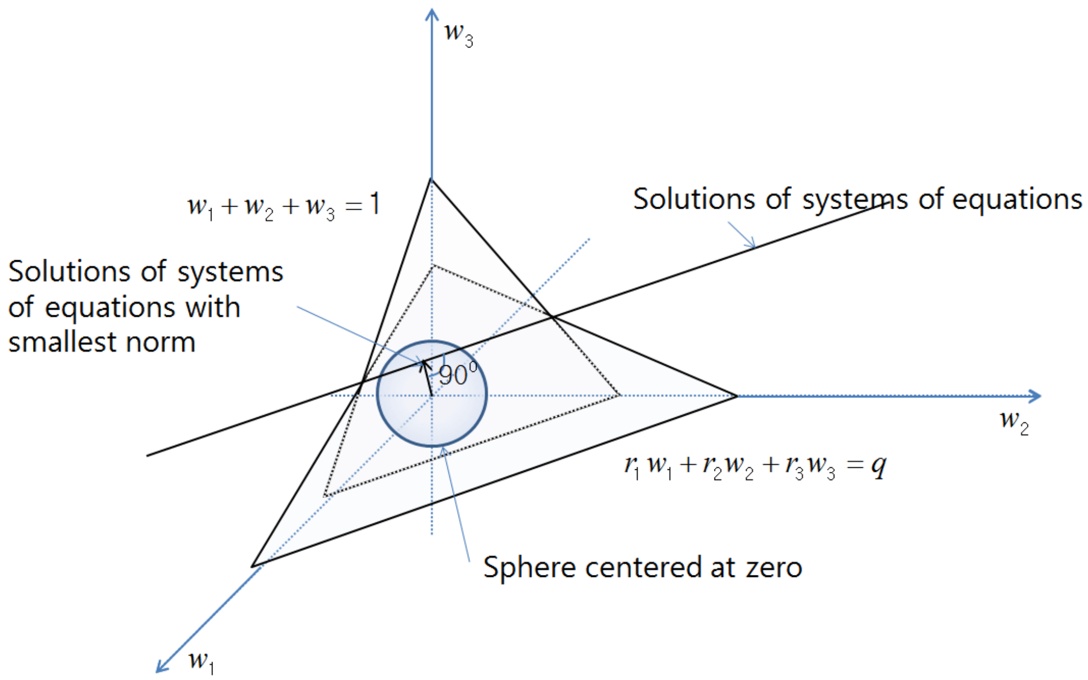

Exhibit 2 shows the idea of new approach when T=1 and N=3. r_1, r_2, r_3 are the obseved return values of three assets at time T=1, respectively. Our approach seeks to find the solution that minimizes the distance from zero to the tangent line which is the intersection of the plane made of $\{ r_1 w_1 + r_2 w_2 + r_3 w_3 = \mathrm{q} \}$ and the plane made of $\{ w_1 + w_2 + w_3 = 1 \}$.

Applying the least squared method to $\boldsymbol{Aw}$=$\boldsymbol{b}$ for a $T \times N$ matrix $\boldsymbol{A}$ implicitly assumes that the number of observations T is larger than the number of unknowns N, which means that $\boldsymbol{Aw}$=$\boldsymbol{b}$ may or may not be consistent, depending on the relative size of observations and unknowns. In other words, there probably will not exist a choice of $\boldsymbol{w}$ that perfectly fits the data $\boldsymbol{b}$, and probably, vector $\boldsymbol{b}$ will not be a combination of the columns of matrix $\boldsymbol{A}$. Therefore, the Markowitz approach to the portfolio optimization problem is valid only when the number of observations is bigger than or equal to the number of assets and when the column rank of matrix $\boldsymbol{R}$ in Equation (4) is full at the same time. Otherwise, the variance-covariance matrix made by $\boldsymbol{R}^T\boldsymbol{R}$ becomes a singular matrix.

Definition: The rank of a matrix $\boldsymbol{A}$ is the dimension of the vector space generated (or spanned) by its columns, which is the same as the dimension of the space spanned by its rows. An alternative definition for the rank is as follows: For an m × n matrix $\boldsymbol{A}$, if f is a linear map from R^n to R^n defined by $f(\boldsymbol{x}) = \boldsymbol{Ax}$, $(\boldsymbol{x} \in \boldsymbol{R}^n)$, then the rank of $\boldsymbol{A}$ is the dimension of the image of f.

Theorem1: Consider the system of equations $\boldsymbol{Ax}=\boldsymbol{b}$. One of the following three possibilities must hold. [4]

(1) If the rank of the augmented matrix [$\boldsymbol{A}$ $\boldsymbol{b}$] is greater than that of $\boldsymbol{A}$, that means $\boldsymbol{b}$ does not belong to column space of $\boldsymbol{A}$, then no solution exists to $\boldsymbol{Ax}=\boldsymbol{b}$.

(2) If the rank of [$\boldsymbol{A}$ $\boldsymbol{b}$] equals that of $\boldsymbol{A}$ and equals the number of unknowns, then the system $\boldsymbol{Ax}=\boldsymbol{b}$ has an exact solution.

(3) If the rank of [$\boldsymbol{A}$ $\boldsymbol{b}$] equals that of $\boldsymbol{A}$ and strictly less than the number of unknowns, then the system $\boldsymbol{Ax}=\boldsymbol{b}$ has infinitely many solutions.

When the system of equations to solve is the case 1 in the above theorem, the most likely solution can be found by getting the least squared solution. For the case 2, the general method to solve the system of equations derives the same solution as with the least squared method. For the case 3 in the above theorem, the least squared method doesn't work and the covariance matrix becomes a singular matrix. There are infinitely many solutions that satisfy $\|\boldsymbol{\varepsilon}\| = 0$.

We use the Singular Value Decomposition (SVD) method to solve a system of linear equations $\boldsymbol{Aw} = \boldsymbol{b}$ as explained in Strang (2006). The SVD is known to be a good method for numerically stable computation. A common use of the SVD is to compute the best fit or least squared solution to a system of linear equations that has no solution or to find the best solution to a system of linear equations that has many solutions.

Let ***S***, ***D***, and ***V*** denote a $\boldsymbol{T} \times \boldsymbol{T}$, $\boldsymbol{T} \times \boldsymbol{N}$, and $\boldsymbol{N} \times \boldsymbol{N}$ matrix, respectively. Then, ***A*** can be factored into:

$$\boldsymbol{A} = \boldsymbol{SDV} = \text{(orthogonal)(diagonal)(orthogonal)} \qquad (7)$$

The columns of $\boldsymbol{S}$ are eigenvectors of $\boldsymbol{AA}^T$, and the rows of $\boldsymbol{V}$ are eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$. The singular values $\sigma_1, \sigma_2; \cdots, \sigma_r$ on the diagonal of $\boldsymbol{D}$ are the square roots of the nonzero eigenvalues of both $\boldsymbol{AA}^T$ and $\boldsymbol{A}^T\boldsymbol{A}$ .Since $\boldsymbol{S}$ and $\boldsymbol{V}$ in Equation (7) are orthogonal matrices, they do not change the length of the other vector when they are multiplied by it.

Let a pseudoinverse of $\boldsymbol{A}$ be denoted by $\boldsymbol{A}^+$ and let $\boldsymbol{D}^+$ be an $\boldsymbol{N} \times \boldsymbol{T}$ matrix that has the reciprocals of the singular values $1/\sigma_1, 1/\sigma_2; \cdots, 1/\sigma_r$ on its diagonal. The pseudoinverse of $\boldsymbol{A}^+$ is $(\boldsymbol{A}^+)^+ = \boldsymbol{A}$. The pseudoinverse $\boldsymbol{A}^+$ is a generalization of the inverse matrix $\boldsymbol{A}^{-1}$ of ***A.*** The pseudoinverse is defined and unique for all matrices whose entries are real or complex numbers.

We can always compute the pseudoinverse $\boldsymbol{A}^+$ using the SVD as follows:

$$A^{+} = V^{T} D^{+} S^{T} \quad (8)$$

Then, we can solve $Aw = b$ by using the pseudoinverse given by Equation (8) as in the following three cases:

(1) If $T = N$, A is a full rank, which implies $A^{+} = A^{-1}$. In this case, $w = A^{+}b = A^{-1}b$.

(2) If $T>N$, $w = A^{+}b$ is the one that minimizes the quantity $\|Aw - b\|$. That is, in this case, there are more constraining equations than variables w so that it is not generally possible to find an exact solution to these equations. Thus, the pseudoinverse given by Equation (8) gives a solution w such that Aw is the "closest" in the least squared sense to the desired vector b.

(3) If $T<N$, there are generally infinite number of solutions, and $w = A^{+}b$ is a particular solution that minimizes the 2-norm of w, denoted by $\|w\|^{2}$ among many solutions. It has the advantage of finding the solution that minimizes $\|w\|^{2}$ because it provides more diversified portfolio weights than the Markowitz solution does.

Theorem2: If A is an $m \times n$ matrix, then A has a singular value decomposition.

The objective of the Markowitz model is to find the portfolio weights that minimize portfolio variance while achieving the target portfolio return on average. However, our approach attempts to find the portfolio that can achieve the target return at each point in time while minimizing the 2-norm of the portfolio weights. Consequently, the in-sample variance of portfolio is always zero in our approach.

The out-of-sample performance of our approach depends on how well the in-sample return vectors $(r_{1t}, r_{2t}, \cdots, r_{Nt})$ at each time t will span the out-of-sample return vectors. If the out-of-sample return vectors were to be identical to the return vectors spanned by the in-sample return vectors at each time *t*, they would satisfy $\sum_{i=1}^{N} w_i r_{it} = \mathrm{q}$. However, since it is impossible to span the unsystematic risk part of the out-of-sample return vectors by the in-sample return vectors, we distinguish between the systematic risk part and the unsystematic risk part and attempt to span the out-of-sample systematic risk part by solely using the in-sample systematic part of return vectors. To impose some structure on the estimation of the systematic part of returns, we assume stock returns are generated by Fama-French five factor model. We denote new model that is based on raw returns as *new method without return structure* and denote new model that uses returns estimated by Fama-French five factor model as *new method with return structure*.

If we decompose $\boldsymbol{A}$ into the systematic risk part $\boldsymbol{A}_*$ and the unsystematic risk part $\boldsymbol{\varepsilon}_{\boldsymbol{A}}$,

$$\boldsymbol{A} = \boldsymbol{A}_* + \boldsymbol{\varepsilon}_{\boldsymbol{A}} \tag{9}$$

Then, the in-sample error should satisfy the following.

$$\boldsymbol{Aw} = \boldsymbol{A}_*\boldsymbol{w} + \boldsymbol{\varepsilon}_A\boldsymbol{w} = \mathrm{q}\boldsymbol{1} + \boldsymbol{\varepsilon}_A\boldsymbol{w} \tag{10}$$

Hence, the in-sample variance $\|\boldsymbol{\varepsilon}_A w\|^2 = w^T \boldsymbol{\varepsilon}_A{}^T \boldsymbol{\varepsilon}_A w$ becomes positive. For the models that explain the systematic risk part $\boldsymbol{A}_*$, we use Fama-French five factor model. [5]

EMPIRICAL RESULTS

To directly compare the results with those of Ledoit and Wolf [2003], we use the same data and portfolio rebalancing strategy as they used. We use the monthly stock returns extracted from the Center for Research in Security Prices (CRSP) from August 1962 to July 1995. We consider common stocks traded on the New York Stock Exchange (NYSE) and the

American Stock Exchange (AMX) with valid CRSP returns for the last 132 months and valid Standard Industrial Classification (SIC) codes.[6] We use S&P 500 as a market index for market model and one month T-bill rate as a risk-free rate. We use the estimates for Fama-French five factors provided by French website.

We use data from August of year t-10 to July of year t as the in-sample period to estimate the mean vector and the covariance matrix of stock returns. Then on the first trading day in August of year t we build a portfolio with minimum variance using identity, pseudoinverse, market model and shrinkage to market estimators for the covariance matrix estimates as Ledoit and Wolf [2003] did and also a portfolio with minimum variance based on our model. We hold this portfolio until the last trading day in July of year t+1, at which time we liquidate it and start the process all over again. Thus, the out-of-sample period goes from August of year t to July of year t+1.

For the purpose of comparison with Ledoit and Wolf [2003], we consider two minimum variance portfolios: the global minimum variance portfolio and the portfolio with minimum variance under the constraint of having 20% expected return.[7] In both cases short sales are allowed. The main quantity of interest is the out-of-sample Sharpe ratio or CEQ return in addition to the out-of-sample mean and standard deviation of this investment strategy over the 23-year period from August 1972 to July 1995. However, since Ledoit and Wolf [2003] are interested in showing what kind of reduction in out-of-sample variance their method yields, they only report the out-of-sample standard deviation of the investment strategy (see Table 1, p. 617 in Ledoit and Wolf [2003]).

Certainty-equivalent (CEQ) return is defined as the risk-free rate that an investor is willing to accept rather than adopting a particular risky return and calculated as follows.

$$\mathrm{CEQ} = \mu_p - \frac{\gamma}{2}\sigma_p^2 \tag{11}$$

, where μ_p =mean of portfolio return, σ_p^2 =variance of portfolio return, and γ =coefficient of relative risk aversion (γ =3 and 5 are assumed).

We consider four covariance matrix estimators proposed in the literature; identity, pseudoinverse, market model and shrinkage to market estimators. The simplest model is to assume that the covariance matrix is a scalar multiple of the identity matrix. When the number of assets *N* exceeds the number of returns *T,* the inverse of the sample covariance does not exist. In this case, replacing the inverse of the sample covariance matrix by the pseudoinverse yields well-defined portfolio weights in the Markowitz model. Market model is the single-index covariance matrix of Sharpe [1963]. Finally, the shrinkage to market estimator that Ledoit and Wolf [2003] recommended is as follows.

$$\hat{\boldsymbol{S}} = \frac{k}{T}\boldsymbol{F} + (1 - \frac{k}{T})\boldsymbol{S} \tag{12}$$

, where $\boldsymbol{F}$ is the single-index covariance matrix, $\boldsymbol{S}$ is the sample covariance matrix and $\frac{k}{T}$ is an optimal shrinkage intensity.

EXHIBIT 3

Out-of-Sample Standard Deviation of Minimum Variance Portfolio

	Unconstrained		Constrained(20%)	
	Standard Deviation	Ledoit -Wolf	Standard Deviation	Ledoit -Wolf
Identity	18.43	17.75	18.54	17.94
Pseudoinverse	12.10	12.37	13.35	13.73

Market model	11.15	12.00	13.12	13.77
Shrinkage to market	8.95	9.55	9.45	10.43
New method without return structure	9.87		10.62	
New method with return structure	10.28		11.46	

Exhibit 3 shows the out-of-sample standard deviation of the minimum variance portfolios for the four covariance matrix estimators and for the two new methods. “Unconstrained” refers to the global minimum variance portfolios while “constrained” refers to the minimum variance portfolios with 20% expected return. Standard deviation is measured out-of-sample at the monthly frequency, annualized through multiplication by $\sqrt{12}$ and expressed in percentage. The columns of Ledoit-Wolf show the results of Table 1 of Ledoit-Wolf [2003], which is almost the same as our results. We can see that the shrinkage to market estimator shows the smallest out-of-sample standard deviation, while new method without return structure is second best and new method with return structure is third best.

EXHIBIT 4

Out-of-Sample Sharpe Ratio and CEQ Return of Minimum Variance Portfolio When In-Sample Period is 10 Years

	Unconstrained				
	STD	Mean	Sharpe Ratio	CEQ(γ=3)	CEQ(γ=5)
Identity	18.43	15.47	0.45	3.18	-0.22
Pseudoinverse	12.10	13.26	0.50	3.87	2.40
Market model	11.15	11.90	0.42	2.84	1.59
Shrinkage to market	8.95	12.56	0.60	4.16	3.35
New method without return structure	9.87	13.71	0.66	5.05	4.07
New method with return structure	10.28	15.00	0.76	6.21	5.16
	Constrained(20%)				
	STD	Mean	Sharpe Ratio	CEQ(γ=3)	CEQ(γ=5)

Identity	18.54	14.41	0.39	2.06	-1.38
Pseudoinverse	13.35	11.58	0.33	1.71	-0.07
Market model	13.12	9.56	0.18	-0.22	-1.95
Shrinkage to market	9.45	11.05	0.41	2.51	1.61
New method without return structure	10.62	11.07	0.37	2.18	1.05
New method with return structure	11.46	12.59	0.47	3.42	2.10

	Constrained(16%)				
	STD	Mean	Sharpe Ratio	CEQ(γ=3)	CEQ(γ=5)
Identity	18.43	15.70	0.46	3.41	0.01
Pseudoinverse	12.27	12.53	0.44	3.07	1.56
Market model	12.14	10.29	0.26	0.88	-0.59
Shrinkage to market	9.14	11.52	0.47	3.07	2.23
New method without return structure	10.10	11.96	0.47	3.23	2.21
New method with return structure	10.71	13.50	0.59	4.58	3.43

Exhibit 4 shows the out-of-sample standard deviation, mean, Sharpe ratio and CEQ return of the minimum variance portfolios for the four covariance matrix estimators and for the two new methods when we maintain the in-sample period as 10 years. We can see that for all cases we consider, new method with return structure shows the highest out-of-sample Sharpe ratio and CEQ return, while new method without return structure and the shrinkage to market estimator are second best or third best, depending on the constraint on expected return. Although the shrinkage to market estimator method shows the smallest out-of-sample standard deviation, it also shows the lowest out-of-sample mean compared to both new methods. Consequently, concerning Sharpe ratio and CEQ return, the shrinkage to market estimator method cannot perform the best. We claim that new model that uses returns estimated by Fama-French five factor model (*new method with return structure*) performs the best.

EXHIBIT 5

Out-of-Sample Sharpe Ratio and CEQ Return of Minimum Variance Portfolio When In-Sample Period is 2 Years

	Unconstrained				
	STD	Mean	Sharpe Ratio	CEQ(γ=3)	CEQ(γ=5)
Identity	19.83	15.30	0.44	2.20	-1.73
Pseudoinverse	19.31	15.15	0.44	2.35	-1.38
Market model	9.96	11.30	0.48	2.62	1.62
Shrinkage to market	9.60	11.58	0.52	2.99	2.07
New method without return structure	12.10	14.38	0.65	4.99	3.52
New method with return structure	12.49	14.31	0.62	4.77	3.21
	Constrained(20%)				
	STD	Mean	Sharpe Ratio	CEQ(γ=3)	CEQ(γ=5)
Identity	18.18	14.65	0.45	2.50	-0.80
Pseudoinverse	18.41	16.08	0.52	3.80	0.41
Market model	9.85	11.54	0.51	2.88	1.91
Shrinkage to market	9.43	11.69	0.54	3.15	2.26
New method without return structure	11.72	13.83	0.62	4.57	3.20
New method with return structure	12.15	13.89	0.60	4.47	3.00

Exhibit 5 shows the out-of-sample standard deviation, mean, Sharpe ratio and CEQ return of the minimum variance portfolios for the four covariance matrix estimators and for the two new methods when we reduce the in-sample period to 2 years.[8] In this case, new method without return structure shows the highest out-of-sample Sharpe ratio and CEQ return, while new method with return structure is second best and the shrinkage to market estimator is third best. We can see that when the in-sample window becomes shorter, the out-of-sample performance of *new method without return structure* that relies on raw returns strengthens.

CONCLUSIONS

We propose a new portfolio optimization approach that does not need the estimation of the covariance matrix and provides solutions to a system of equations using the Singular Value Decomposition (SVD) method with the additional constraint that the 2-norm of the portfolio weights is minimized. By imposing some structure on stock returns, we show our new approach can enhance the out-of-sample performance of portfolio further.

To directly compare the results with those of Ledoit and Wolf [2003], we use the same data and portfolio rebalancing strategy as they used. When we assume the in-sample window is 10 years, new method with return structure shows the highest out-of-sample Sharpe ratio and CEQ return, while new method without return structure and the shrinkage to market estimator are second best or third best, depending on the constraint on expected return. Although the shrinkage to market estimator method shows the smallest out-of-sample standard deviation, it also shows the lowest out-of-sample mean compared to both new methods. Consequently, concerning Sharpe ratio and CEQ return, the shrinkage to market estimator method cannot perform the best. Furthermore, we find that when the in-sample window becomes shorter and less than 4 years, the out-of-sample performance of new method without return structure that relies on raw returns strengthens. In this case, new method without return structure shows the highest out-of-sample Sharpe ratio and CEQ return, while new method with return structure is second best and the shrinkage to market estimator is third best.

ENDNOTES

[1]DeMiguel et al. [2008] suggests a new approach for determining the optimal portfolio weights in the presence of estimation error by solving the traditional minimum variance problem with the additional constraint that the norm of the portfolio weight vector be smaller than a given threshold.

[2] The bold small letters in the equation denote the column vector notation, and the bold capital letters denote the matrix notation.

[3] Ignoring (1/(T-1)) does not affect the problem.

[4] See p.146 of Ben and Daniel [1987].

[5] We did the same analysis using the market model and Fama-French three factor model which was omitted.

[6] Although Ledoit and Wolf [2003] used 120 months as a non-missing period for the sample, we use 132 months to eliminate the possibility of delisted companies among the sample which could change the composition of portfolio during the out-of-sample period.

[7] In order to see the effect of the change in expected return on the out-of-sample performance of Sharpe ratio and CEQ return, we consider the constraint of 16 % as well.

[8] Although we do not report in the paper, we did the same analysis for different in-sample windows and found that for the in-sample windows less than 4 years, new method without return structure performs the best and for the in-sample windows more than or equal to 5 years, new method with return structure performs the best.

ACKNOWLEDGEMENTS

We wish to thank Byung-wook Choi for his valuable feedback and Matlab programming. This paper was financially supported by Konkuk University in 2013.

REFERENCES

DeMiguel, V., L. Garlappi, F. Nogales, R. Uppal, “A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms.” Management Science, Vol. 55, No. 5 (2009), pp. 798-812.

Fama, E., K. French, “A Five-Factor Asset Pricing Model.” Journal of Financial Economics, Vol. 116, No. 1 (2015), pp. 1-22.

Frost, P., J. Savarino, “For Better Performance: Constrain Portfolio Weights.” Journal of Portfolio Management, 15 (1998), pp. 29-34.

——. “An Empirical Bayes Approach to Efficient Portfolio Selection.” Journal of Financial and Quantitative Analysis, 21(1986), pp.293-305.

Herold, U., R. Maurer, “Portfolio Choice and Estimation Risk. A Comparison of Bayesian to Heuristic Approaches.” Astin Bulletin, 36 (2006), pp. 135-160.

Jagannathan, R., T. Ma, “Risk Reduction in Large Portfolios: Why imposing the Wrong Constraints Helps.” Journal of Finance, 58 (2003), pp. 1651-1653.

Jorion, P., “Portfolio Optimization in Practice.” Financial Analysts Journal, 48 (1992), pp. 68-74.

Ledoit, O., M. Wolf, “Improved Estimation of the Covariance Matrix of Stock Returns with an Application to Portfolio Selection.” Journal of Empirical Finance, 10 (2003) pp. 603-621.

Litterman, R., Modern Investment Management: An Equilibrium Approach, Wiley, New York, 2003.

Markowitz, H., “Portfolio Selection.” Journal of Finance, 7 (1952), pp. 77-91.

Michaud, R. O., “Efficient Asset Management.”Harvard Business Review Press, Boston, 1998.

Noble B., J. Daniel, Applied Linear Algebra, 3rd ed. Pearson, 1987.

Pappas, D., K. Kiriakopoulos, G. Kaimakamis, “Optimal Portfolio Selection with Singular Covariance Matrix.” International Mathematical Forum, 5 (2010), pp. 2305 – 2318.

Sengupta, J. K., “Optimal Portfolio Choice in the Singular Case.” International Journal of Systems Science, 14 (1983), pp. 995-1012.

Sharpe, W., “A Simplified Model for Portfolio Analysis.” Management Science Vol. 9, No. 2 (1963), pp. 277-293.

Strang, G., Linear Algebra and Its Applications, 4th ed. Cengage Learning, 2006.