International Mathematical Forum, 5, 2010, no. 47, 2305 - 2318

Optimal Portfolio Selection with Singular Covariance Matrix

Dimitrios Pappas

Department of Statistics, Athens University of Economics and Business
76 Patission Str, 10434, Athens, Greece
pappdimitris@gmail.com

Konstantinos Kiriakopoulos

Proton Bank, Capital Markets Director
Lagoumitzi 22, 17671 Athens Greece
k_kiriak@otenet.gr

George Kaimakamis

Hellenic Army Academy, Vari 16673, Athens Greece
gmiamis@gmail.com

Abstract

In this paper we use the Moore-Penrose inverse in the case of a close to singular and ill-conditioned, or singular variance-covariance matrix, in the classic Portfolio Selection Problem. In this way the possible singularity of the variance-covariance matrix is tackled in an efficient way so that the various application of the Problem to benefit from the numerical tractability of the Moore-Penrose inverse.

Mathematics Subject Classification: 15A09, 91B02, 91B28

Keywords: Covariance matrix, Portfolio selection, Moore-Penrose inverse matrix, optimal portfolio positions, ill-conditioned matrix

1 Introduction

The objective of this paper is to present an application of the generalized inverse matrix to Portfolio Selection Problem (Markowitz (1952),[10]) either in its static or in its multi objective dynamic form (Audet, Savard, (2007

[14]). It is well known fact that a central issue in empirical Finance with implications for portfolio selection is the covariance matrix[1] of stock returns. These implications cover the whole spectrum of finance and can be easily applied to specific areas such as the bond market (Korn O., Koziol C.(2006 [7]). Usually the covariance matrix is estimated using modern statistical techniques (Olivier Ledoit - Michael Wolf 2003 [8]). This covariance matrix is used to calculate the portfolio weights . The sample variance covariance matrix is seldom used because it imposes too little structure, since either it is non-singular or numerically ill-conditioned. Various attempts to tackle this problem appear in the literature such as Buser (1973) [3] , or the "corrected" version of Peter J.Ryan-Jean Lefoll (1981 [9]). Furthemore criteria based on the mean-square error can produce satisfactory solutions (Jati K. Sengupta(1983) [13]) but the use of generalized inverse is more broad and can be applied easily in numerical applications of a large magnitude.

When the number of stocks N is of the same order of magnitude as the number of returns per stock T, the total number of parameters to estimate is of the same order as the total size of the data set. When N is larger than T the covariance matrix is always singular. In this case the problem is that we need the inverse of the covariance matrix and it does not exist. To get around to this problem we can use the generalized inverse or Moore-Penrose inverse. Especially if we replace the inverse of the sample covariance matrix by the pseudo-inverse we can define the portfolio weights w_i. Consider a universe of N stocks whose returns are distributed with mean vector μ and covariance matrix Σ. The problem of portfolio selection is, as defined by Markowitz in [10] and [11]:

$$\min w'\Sigma w$$

subject to $w'1 = 1$ and $w'\mu = q$, where 1 denotes a conformable vector of ones and q is the expected rate of return that is required on the portfolio. Negative elements of w denote short positions.
The well-known solution is:

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

where $A = 1'\Sigma^{-1}1, \quad B = 1'\Sigma^{-1}\mu, \quad C = \mu'\Sigma^{-1}\mu$

This equation shows that optimal portfolio weights depend on the inverse of the covariance matrix. This sometimes causes difficulty if the covariance matrix estimator is not invertible, close to singular or numerically ill-conditioned, which means that inverting it amplifies estimation error tremendously. One possible trick to get around this problem is to use the pseudo-inverse, also called generalized inverse or Moore-Penrose inverse. Replacing the inverse of

[1]The terms variance,covariance and variance-covariance matrix are used interchangably throughout this paper

the sample covariance matrix by the pseudo-inverse into equation (1) yields well-defined portfolio weights. In practice, the covariance matrix is estimated from historical data available up to a given date, optimal portfolio weights are computed from this estimate, the portfolio is formed on that date and held until the next rebalancing occurs. The performance of a covariance matrix estimator is measured by the variance of this optimal portfolio after it is formed. It is a measure of out-of-sample performance, or of predictive ability. An estimator that overfits in-sample data can turn out to work very poorly for portfolio selection, which is why imposing some structure is beneficial.

The notion of the generalized inverse of a (square or rectangular) matrix was first introduced by H. Moore in 1920, and again by R. Penrose in 1955. These two definitions are equivalent, and the generalised inverse of a matrix is also called the Moore- Penrose inverse.

Let A be a $r \times m$ real matrix. Equations of the form $Ax = b$ occur in many pure and applied problems. When A is a singular square matrix, then its unique generalized inverse $A^\dagger$ (known as the Moore- Penrose inverse) is defined. Penrose showed that there is a unique matrix satisfying the four Penrose equations (defined below) and is called the generalized inverse of A, noted by $A^\dagger$.

Apart from the Moore-Penrose inverse, other types of generalized inverses are also in use, such as the Drazin inverse , the Tseng inverse and the Group inverse.

Standard references on generalized inverses are the books of Ben-Israel and Greville [1], Campbell and Meyer [4] and Groetsch [5].

2 Preliminary Notes

We shall denote by $\mathcal{R}^{r\times m}$ the linear space of all $r \times m$ real matrices. For $A \in \mathcal{R}^{r\times m}$, $R(A)$ will denote the range of A and $N(A)$ the kernel of A. The generalized inverse $A^\dagger$ (known as the Moore- Penrose inverse) is the unique matrix that satisfies the following four Penrose equations:

$$AA^\dagger = (AA^\dagger)^*, \qquad A^\dagger A = (A^\dagger A)^*, \qquad AA^\dagger A = A, \qquad A^\dagger AA^\dagger = A^\dagger,$$

where A^* denotes the transpose matrix of A.

It is easy to see that $\mathcal{R}(A^\dagger) = \mathcal{N}(A)^\perp$, $AA^\dagger$ is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{R}(A)$ and that $A^\dagger A$ is the orthogonal projection of $\mathcal{H}$ onto $\mathcal{N}(A)^\perp$. It is well known that $\mathcal{R}(A^\dagger) = \mathcal{R}(A^*)$.

Let us consider the equation $Ax = b, A \in \mathcal{R}^{r\times m}, b \in \mathcal{R}^r$. If A is a square matrix and invertible, then the above equation is in general easy to solve. If A is an arbitrary matrix, then there may be none, one or an infinite number of solutions, depending on whether $b \in R(A)$ or not, and on the rank of A. But if $b \notin R(A)$, then the equation has no solution. Therefore, another point of

view of this problem is the following: instead of trying to solve the equation $\|Ax-b\|=0$, we are looking for a vector u that minimizes the norm $\|Ax-b\|$. Note that the vector u is unique. Since we are interested in the distance between Ax and b, it is natural to make use of $\|A\|_2$ norm. So, in this case we consider the equation $Ax = P_{R(A)}b$, where $P_{R(A)}$ is the orthogonal projection on $\mathcal{R}(A)$.
The following two Theorems can be found in [5].

Theorem 2.1 *Let $A \in \mathcal{R}^{r\times m}$ and $b \in \mathcal{R}^r, b \notin R(A)$. Then, for $u \in \mathcal{R}^m$, the following are equivalent:*

(i) $Au = P_{R(A)}b$

(ii) $\|Au-b\| \leq \| Ax-b\|, \forall x \in \mathcal{R}^m$

(iii) $A^*Au = A^*b$

Let $\mathcal{B} = \{u \in \mathcal{R}^m | A^*Au = A^*b\}$. This set of solutions is closed and convex, therefore, it has a unique vector with minimal norm. Note that, in the literature [5], $\mathcal{B}$ is known as the set of the generalized solutions.

Theorem 2.2 *Let $A \in \mathcal{R}^{r\times m}$ and $b \in \mathcal{R}^r, b \notin R(A)$, and the equation $Ax = b$. Then, if $A^\dagger$ is the generalized inverse of A, we have that $A^\dagger b = u$, where u is the minimal norm solution defined above.*

We shall make use of this property of the Moore-Penrose Inverse to minimize the risk in portfolio selection. Since the covariance matrix is self-adjoint ($\Sigma = \Sigma^*$), it is well known that $\Sigma^\dagger = \Sigma^{\dagger *}$.
An interesting property of self adjoint matrices, is that their Moore-Penrose inverse coincides with two other types of generalized inverses, the Drazin inverse and the Group inverse.

3 Main Results

4 The Moore-Penrose approach: Computations and comparison results

In this work, we will consider 3 cases. A close to singular covariance matrix, a close to singular but numerically ill-conditioned covariance matrix, and a singular covariance matrix. As we will present, when the covariance matrix is close to singular, then, if we replace Σ^{-1} with $\Sigma^\dagger$, the results coincide.
In the case of a close to singular, but ill-conditioned covariance matrix,(a large

condition number) the use of the Moore-Penrose Inverse $\Sigma^\dagger$ gives marginally better results than Σ^{-1}. Last,when the matrix Σ is singular, then we propose $\Sigma^\dagger$ as a candidate for the minimizer, in order to achieve the optimal portfolio positions.

The proof of Markowitz's problem,is performed using the standard Langrage method. The conditions for the Lagrangian give the equation

$$\Sigma w = -\frac{1}{2}(\lambda_1 \mu + \lambda_2 1) \tag{2}$$

When the covariance matrix Σ is singular, then from Proposition 2.2, the minimum norm solution (i.e the optimal portfolio positions) of this system of equations is

$$\hat{w} = -\frac{1}{2}(\lambda_1 \Sigma^\dagger \mu + \lambda_2 \Sigma^\dagger 1) \tag{3}$$

The uniqueness of the solution is due to the uniqueness of the Moore-Penrose Inverse.
Using this vector, we have that the optimum portfolio selection is given by

$$\hat{w} = \frac{C - qB}{AC - B^2}\Sigma^\dagger 1 + \frac{qA - B}{AC - B^2}\Sigma^\dagger \mu \tag{4}$$

where $A = 1'\Sigma^\dagger 1, \quad B = 1'\Sigma^\dagger \mu, \quad C = \mu'\Sigma^\dagger \mu$

In the present work, all the computations have been performed using Matlab programming language.
All the data used in this paper are provided by Bloomberg.

4.1 A close to singular covariance matrix

This study uses a portfolio consisting of 43 stocks, 20 from FTSE-20 Index, and 23 from DAX-30 Index[2]. The period under examination extends from 1/1/2002 to 31/12/2008, with total of 1747 observations. The covariance matrix Σ of their returns is a 43×43 matrix, very close to singular [3]($\det \Sigma = 1.9272 * 10^{-50}$) and with condition number 126.65 . A large condition number indicates a nearly singular matrix, and is an index of the numerical stability of the solution, in our case the inverse Σ^{-1}.
The difference between the norm of Σ and $\Sigma^\dagger$ is of the magnitude of 10^{-13}
We have used the following methodology: In equation (1) we make use of its

[2]The FTSE-20 Index is the market capitalization index with the 20 most heavy capitilized stocks in the Athens Stock Exchange. Also Dax-30 is the index with the 30 most capitalized stocks in Frankfurt Stock Exchange. We have used for the DAX-30 only 23 stocks due to lack of data for the rest 7 ones

[3]standard hypothesis about the lognormality of the daily returns has been assumed. Also it is assumed that each year contains 250 trading days

generalized inverse $\Sigma^{\dagger}$ instead of Σ^{-1} and then compare the results given by both matrices.
Our first step is to make use of the generalized inverse $\Sigma^{\dagger}$ and compute the weights of the stocks of the portfolio depending on q, where q is the expected rate of return for this specific portfolio. We are using the 3-month euribor rate[4] $q = 2,8488\%$ as a starting expected rate of return, and make a graph to examine the relation between the expected returns and the variance of the selected portfolios. The expected returns take values from the interval $[2,8488\% - 10\%]$, using as step $i = 0.005$. As expected, as the $q's$ are growing, the risk is getting bigger.
As we can see in Figure 1 , for expected return $q = 6,15\%$ the variance of the portfolio is minimized and has the value $1,68\%$.

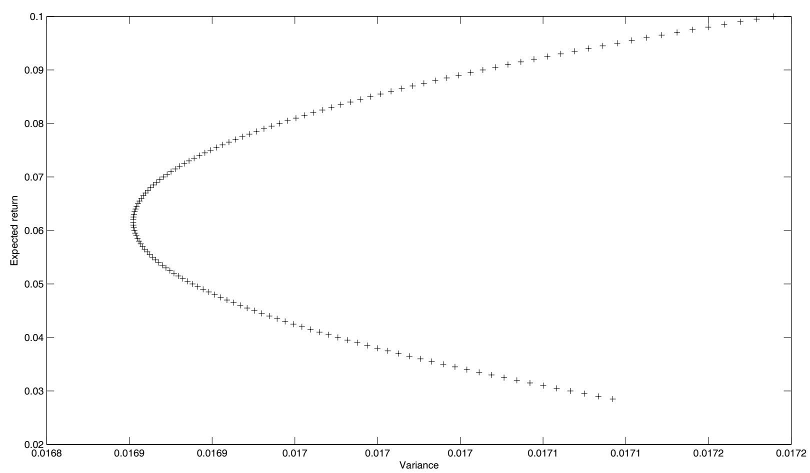

Figure 1: Expected returns vs variance of the portfolio

The next step is to use the inverse Σ^{-1} for all the computations as above, and compare the results. As expected the two results coincide.
Finally we examine the portfolio weights under the two matrices. We find upper bound for the given w_i 's in the two cases and try to see which one has the minimum norm, a measure indicating the difficulty of the implementation of short positions, especially in illiquid markets or in markets where shorts positionning is restrictive. Using (1) we get:

$$\|w\| \leq \| \frac{\Sigma^{-1}}{AC - B^2} \| \left(\|C - qB\| + \|qA - B\| \right)$$

[4]euribor stands for the european interbank offered rate and its the most common reference rate between banks.

which means that:

$$\|w\| \leq \| \frac{\Sigma^{-1}}{AC - B^2} \| \left(q(\|B\| + \|\mu A\|) + \|C\| + \|\mu B\|\right)$$

The results that we get for the two cases are similar although for the generalized inverse are marginally better[5]. This enables to infer that for large portfolio applications, the use of generalized inverse is preferred since it produces smaller norm of the portfolio weights, indicating smaller transaction costs, something which is important for frequent trading.

We can see that

$$\|w\| \leq M_1$$

By replacing in the above equation $\Sigma^{\dagger}$ where Σ^{-1} is and following the same procedure, we have that

$$\|w\| \leq M_2$$

(see Appendix)
A natural question to ask is: In which cases the norm given by the generalized inverse is lower than the one given by the inverse ? In other words, when is the upper bound M_2 lower that M_1 ?
By solving the inequality $M_2 \leq M_1$, we see that when the expected return $q \leq \hat{q} = 5.978\%$, then the portfolio weights given by the generalized inverse have lower norm bounds than the ones given by the inverse.

Short positions are another tool for judging which method provides us the most easy to trade portfolio (e.g [2]), in the sense that it requires the least amount of short selling (which is essentially the sum of all negative portfolio weights). This is due to the fact that short selling is hard to carry out in practice for most investors. We can see in figure 2 a graph of the short selling vs the variance and the expected returns.
The two methods show similar values of the minimum short interest.
(see Appendix).

4.2 An ill-conditioned close to singular covariance matrix

In this study, we use a covariance matrix of a portfolio consisting of 7 intraday exchange rates with 1 minute interval. The exchange rates are the following: Euro Swiss Franc(EURCHF), Euro/ British Pound(EURGBP), Euro / Japanese Yen(EURJPY), Euro / US Dollar(EURUSD), British Pound/ US

[5]The upper bound for w_i using the generalized inverse is smaller in the magnitude of 10^{-13}.

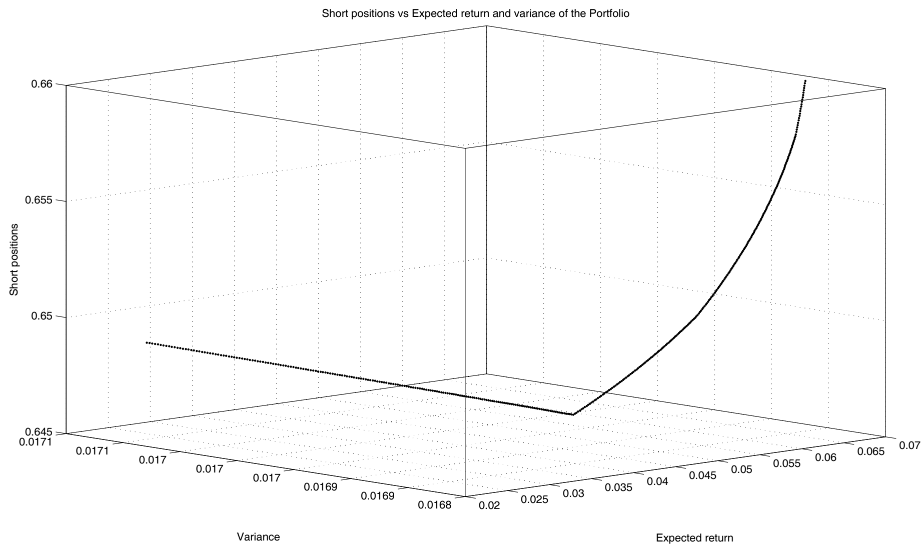

Figure 2: Short positions of the Portfolio

Dollar(GBPUSD), US Dollar / Swiss Franc(USDCHF) and US Dollar/Japanese Yen(USDJPY). Our data comes from 21 October 2008, time 00.00 to 22 October 2008, time 23.59, a total of 2880 intervals. As we can see in the next figure, the determinant of the corresponding covariance matrices (using a 5-minutes interval) is very often equal to zero.

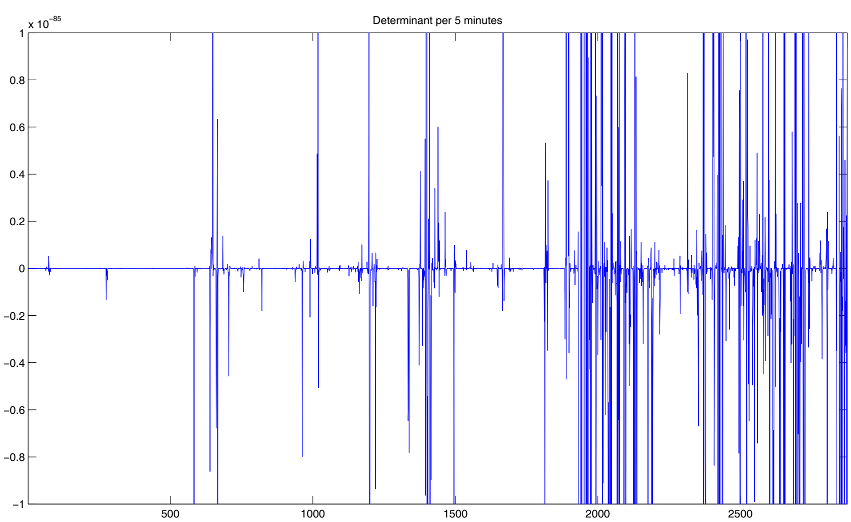

Figure 3: Time intervals vs the Determinant of the covariance matrix

For example, by using data from 22 October 2008, time 23.53 to 24.00 and supposing we want to reevaluate our portfolio every 7 minutes (every 7 intervals), the covariance matrix is very close to singular.

The covariance matrix used in our case (a 7×7 matrix) has determinant very close to zero ($\det \Sigma = 1.76 * 10^{-77}$) and with very large condition number ($5.9323 * 10^{16}$). Therefore the numerical computation of its inverse has large estimation error.
In the following figures, the results concerning the Moore-Penrose inverse will appear in red colour, and those concerning the Inverse in black.
In Figure 4 we present the graph using the Inverse and the Moore-Penrose inverse of the covariance matrix. We can see that the results are much more accurate in the case of the Moore-Penrose Inverse, especially when the expected return is more than 3%. Also the larger the expected return, the better the results (expressed through lower variance) in the generalized inverse case.

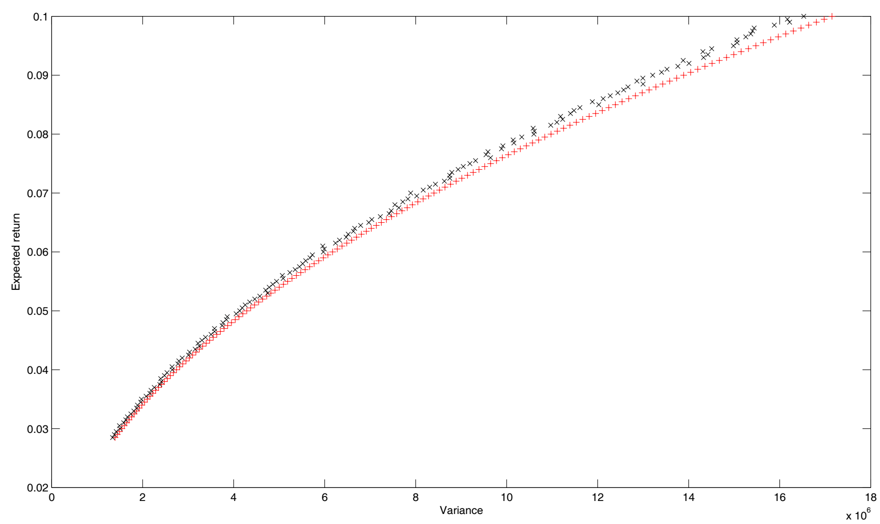

Figure 4: Expected returns vs Variance . Inverse (x) and Moore-Penrose Inverse (+)

We can also examine which method provides us the most attractive portfolio, like in the previous case. We will examine the short selling using the two methods: As we can see in figure 5, in this case the two methods do not show similar values of the minimum short positions. (See Appendix for details)

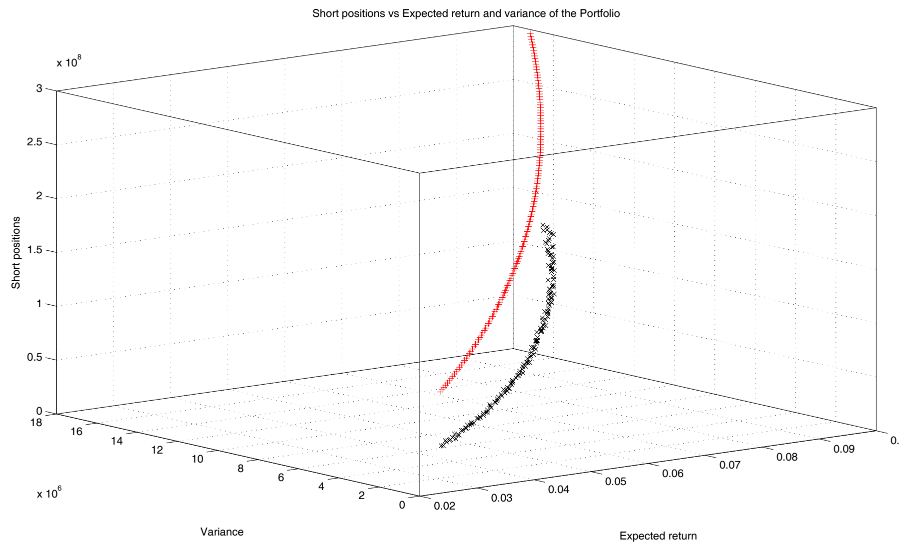

Figure 5: Short positions. Inverse (x) and Moore-Penrose Inverse (+)

The Sharpe ratio [6]is another index for comparing the results. The Sharpe ratio is used to characterize how well the return of an asset compensates the investor for the risk taken. When comparing two assets each with the expected return q against the same benchmark, the asset with the higher Sharpe ratio gives more return per unit of risk.
We can see in figure 6 that the results are almost equal, with a difference of $2 * 10^{-7}$ at the most.

4.3 A singular covariance matrix

In this case we will use the same example as in the previous case, but the difference is that the data used are from 21 October 2008 , time 22.57 until 23.00.(3 intervals) The covariance matrix is singular (det $\Sigma = 0$).
In this case, as presented above, the use of the Moore-Penrose inverse is necessary for finding optimal portfolio positions.
In figures 7 and 8, we present the results using the Moore-Penrose Inverse: (Variance vs expected returns, short portfolio positions vs expected returns and Sharpe ratio vs expected returns).

[6]It is formally defined as the excess return attained per unit of portfolio risk; $S = \frac{(q-r_f)}{\sqrt{w'\Sigma w}}$, where in our case the risk free rate is assumed to be $r_f = 2,85\%$

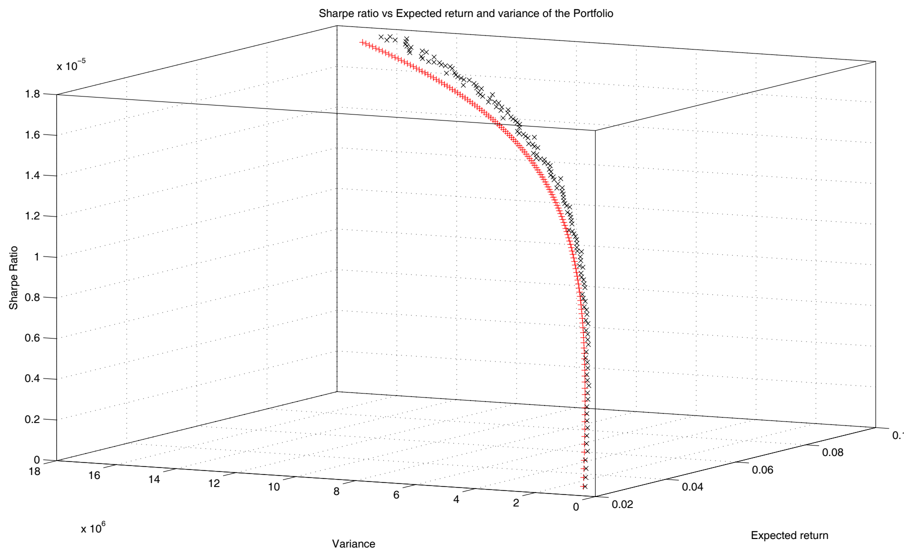

Figure 6: Sharpe ratio vs Variance and expected return. Inverse (x) and M-P Inverse(+)

5 Elapsed Computational time

There are several methods for computing the Moore-Penrose inverse matrix (cf. [1]). One of the most commonly used methods is the Singular Value Decomposition (SVD) method. This method is very accurate but also time-intensive since it requires a large amount of computational resources, especially in the case of large matrices. In a recent work, D. Pappas and V. Katsikis [6] presented a new algorithm for the computation of the Moore-Penrose inverse matrix. This method is very fast and has the same accuracy as Matlab's function "pinv". Using this method, the elapsed time for computing the generalized inverse of the covariance matrix in the first case of our study (a 43×43 matrix) using this method was 0.028102 seconds. By computing the same matrix using Matlab's built- in function, the elapsed time was 0.099287 seconds. As we can see, this method is about 3.5 times faster, which is crucial in many large scale on-line applications and it can be very useful for mechanical on-line trading systems.

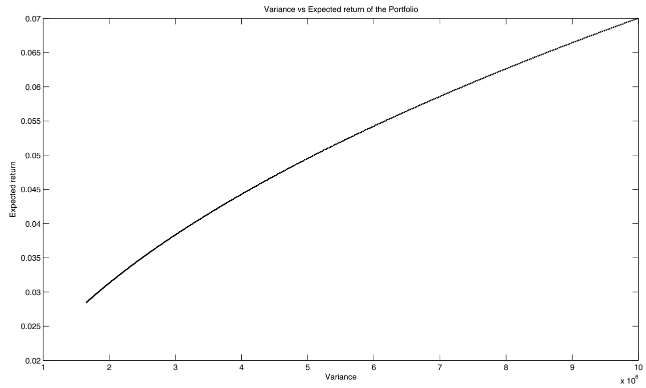

Figure 7: Expected return vs Variance

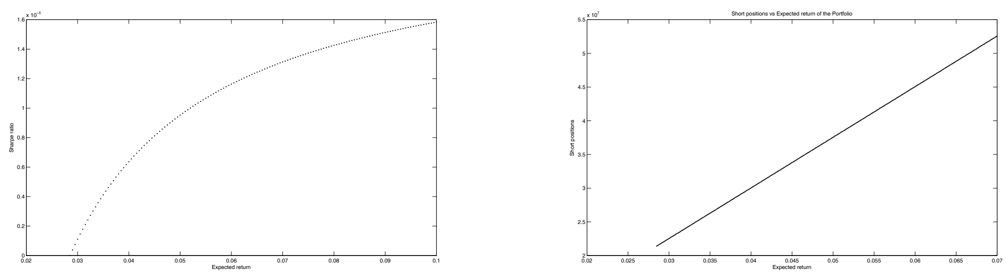

Figure 8: Sharpe ratio vs Expected return (a), Short positions vs Expected return (b)

6 Conclusions

We have seen that when the variance matrix is singular the use of the generalized inverse results in optimal portfolio weights in the classical Markowitz problem. Moreover as it is shown from the above computations, when the covariance matrix is close to singular, the use of the generalized inverse matrix instead of the classical inverse matrix produces almost equal results.
By replacing Σ^{-1} by the generalized inverse $\Sigma^{\dagger}$ does not affect the portfolio results , but the use of generalized inverse gives lower upper bound for the portfolio weights up to a certain number of expected returns $\hat{q}$.
In case when the covariance matrix is numerically ill- conditioned, the results given using the generalized inverse instead of the inverse are marginally better. The same conclusions can be derived concerning the short interest and the Sharpe ratio of the proposed portfolios.
So in any case the use of generalized inverse is recommended for large portfolio applications when the covariance matrix is close to singular, ill conditioned or

singular.

7 Appendix

1. First Case: (Subsection 3.1)
   After computations, we have that

   $$\|w\| \leq M_1 = 12.901397238349661 * q + 1.801079615018838$$

   by using the inverse Σ^{-1}.
   By replacing it by $\Sigma^{\dagger}$ and following the same procedure, we have that

   $$\|w\| \leq M_2 = 12.901397238349702 * q + 1.801079615018835.$$

   Short Positions: Using the generalized inverse matrix , 64.6722270399304% of all portfolio positions must be shorted, when the expected return is $q = 4.82\%$, whereas using the inverse matrix the percentage 64.6722270399305% of short positions with the same expected return. The variance of the minimum short positions portfolio in both cases is equal to 1.69%.

2. Second case: (Subsection 3.2)
   In this case, the amount of short positions is very large, in the magnitude of 10^7. The difference between the two methods is in the magnitude of $1 * 10^7$.

ACKNOWLEDGEMENTS. The authors would like to thank Professor A. Yannacopoulos for his very helpful suggestions and PhD student A. Petralias for many interesting conversations on this subject.

References

[1] A. Ben-Israel and T. N. E. Grenville, *Generalized Inverses: Theory and Applications*, Springer- Verlag, Berlin 2002.

[2] C. Bengtsson, J. Holst: On portfolio selection: Improved Covariance matrix estimation for Swedish Asset returns, Technical Report, Department of Economics, Lund University, Sweden, Oct. 2002.

[3] S,A. Buser: Mean-Variance Portfolio Selection with Either a Singular or Non-Singular Variance-Covariance Matrix, Journal of Financial and Quantitative Analysis. **81I**, (May/June 1973), 637-654.

[4] S. L. Campbell and C. D. Meyer *Generalized inverses of Linear Transformations*, Dover Publ. New York 1979.

[5] C. Groetsch : Generalized inverses of linear operators, Marcel Dekker 1977

[6] V. Katsikis and D. Pappas, "Fast computing of the Moore- Penrose Inverse matrix", Electronic Journal of Linear Algebra,**17** (Nov 2008), 637- 650.

[7] Korn O.-Koziol C.:Bond Portfolio Optrimization A Risk-Return Approach, The Journal of Fixed Income. March 2006

[8] O. Ledoit - M. Wolf: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, Journal of Empirical Finance 10 (2003), 603- 621

[9] P.J.R. Lefoll: A Comment on Mean-Variance Selection with either a singular or a non-singular variance-covariance Matrix, Journal of Financial and Quantitative Analysis. Vol.XVI, No 3 (September 1981), 389-395

[10] H.Markowitz: Portfolio Selection, The Journal of Finance, **7**, No.1. (Mar. 1952), 77-91.

[11] H.Markowitz: Portfolio Selection, Cowles Foundation Monograph N.16. New York John Wiley and Sons Inc., 1959

[12] A.P. Martins: Portfolio Selection-A Technical Note. Seminar of Faculdade de Ciencias Economicas e Empresariais, Universidade Catolica Portuguesa (2005)

[13] J. K. Sengupta, Optimal Portfolio Choice in the Singular Case, International Systems Science, **14**, No 8 (1983), 995-1012.

[14] W. Zgal ,C. Audet ,G. Savard : A New Multi Objective Approach for the Portfolio Selection Problem with Skewness, Les Cahiers du Gerard and Department de mathematique et de genie industriel, G-2007-86.

Received: March, 2010