Optimal mean-variance portfolio selection using Cauchy–Schwarz maximization

Hsin-Hung Chen\textsuperscript{a}, Hsien-Tang Tsai\textsuperscript{b,\textasteriskcentered} and Dennis K. J. Lin\textsuperscript{c}

\textsuperscript{a}Department of Business Administration, Cheng Shiu University, Kaohsiung County, Taiwan
\textsuperscript{b}Department of Business Management, National Sun Yat-Sen University, Kaohsiung, Taiwan
\textsuperscript{c}Department of Statistics, Pennsylvania State University, University Park, USA

Fund managers highly prioritize selecting portfolios with a high Sharpe ratio. Traditionally, this task can be achieved by revising the objective function of the Markowitz mean-variance portfolio model and then resolving quadratic programming problems to obtain the maximum Sharpe ratio portfolio. This study presents a closed-form solution for the optimal Sharpe ratio portfolio by applying Cauchy–Schwarz maximization and the concept of Kuhn–Tucker conditions. An empirical example is used to demonstrate the efficiency and effectiveness of the proposed algorithms. Moreover, the proposed algorithms can also be used to obtain the optimal portfolio containing large numbers of securities, which is not possible, or at least is complicated via traditional quadratic programming approaches.

I. Introduction

The mean-variance portfolio model, proposed by Markowitz (1952, 1959), has served as the guide for most subsequent asset allocation models. Markowitz focused on portfolios rather than individual securities. The innovation of Markowitz was followed in the 1960s by the Capital Asset Pricing Model (CAPM), which was articulated most notably by William F. Sharpe. In 1990, Markowitz, Sharpe and Miller were awarded Nobel Prizes owing to their achievements in improving and popularizing the mean-variance portfolio model. However, the mean-variance model has also been criticized. For example, Borch (1969) and Feldstein (1969) indicated that the mean-variance framework only leads to optimal decisions if utility functions are quadratic or investment returns are jointly elliptically (or spherically) distributed. Therefore, Bawa and Lindenberg (1977) proposed a portfolio model known as the Mean-Lower Partial Moment (MLPM) portfolio framework based on the concept of downside risk.

During the 1990s, the popularity of downside risk among investors grew and the mean return-downside risk portfolio selection model seemed to be superior to the mean-variance framework (Grootveld and Hallerbach, 1999). However, Grootveld and Hallerbach (1999) examined the differences and similarities between the variance and downside risk measures, and published an important article. Contradicting common beliefs, their study...
demonstrated that few members of the large family of downside risk measures possess better theoretical properties within a return-risk framework than does variance. Moreover, the implementation of mean-downside risk portfolio models is much more tedious since there are no shortcuts in computing portfolio risk (Grootveld and Hallerbach, 1999).

Consequently, the mean-variance model has remained the most important portfolio framework during recent years. Numerous scholars have successfully continued to study and revise mean-variance model, such as Huang and Litzenberger (1988), Elton and Gruber (1995), Elliott and Kop (1999), Jorion (2003), Mercurio and Torricelli (2003), Prakash et al. (2003), Ehrrott et al. (2004), Ambachtsheer (2005), Campbell and Viceira (2005), Aquino (2006) and Ulucan (2007). The most recent research by Ulucan (2007) investigated optimal holding period (investment horizon) for the classical mean-variance portfolio model. The historical transaction records of Istanbul Stock Exchange ISE-100 index stocks and Athens Stock Exchange FTSE-40 index stocks data were used in empirical analysis. The results of that research showed that portfolio returns with varying holding periods had a convex structure with an optimal holding period.

The mean-variance portfolio model defines risk in terms of the possible variation of expected portfolio returns. Moreover, the ‘efficient frontier’ is the set of the portfolios generated by the Markowitz mean-variance portfolio model with the highest achievable expected returns for given SDs or the lowest achievable SDs for given expected returns. However, the portfolios on the efficient frontier generally exhibit the higher return with higher risk characteristic, often creating a dilemma for portfolio management decision makers. Although Markowitz (1959) proposed applying the expected utility maximum method to determine the optimal portfolio for investors, optimal portfolio selection using the mean-variance model has been a problem and received considerable discussion (see e.g., Borch, 1969; Feldstein, 1969; Levy and Markowitz, 1979; Kroll et al., 1984).

On the other hand, Sharpe (1966) introduced the Sharpe ratio for the performance of mutual funds and portfolio selection. The Sharpe ratio is built on Markowitz’s mean-variance paradigm, which assumes that the mean and SD of the distribution of one-period returns are sufficient statistics for evaluating the prospects of an investment portfolio (Sharpe, 1994). Since Sharpe introduced the Sharpe ratio, most financial institutions have used it to evaluate the performance of mutual funds and select portfolios. Although various measures have been proposed for evaluating the performance of portfolios (see e.g., Dowd, 2000; Campbell et al., 2001), the Sharpe ratio is still a major index to measure the performance of mutual funds. Moreover, this ratio can be used to select the optimal portfolio on the efficient frontier generated by the Markowitz mean-variance model because it considers both the mean and the SD of the portfolio returns.

Specifically, fund managers can revise the objective function of Markowitz mean-variance model and then apply quadratic programming techniques to obtain the maximum Sharpe ratio portfolio. However, the number of company stocks is increasing in stock exchange markets. For example, the New York Stock Exchange (NYSE) market already has more than 2800 company stocks, while the National Association of Securities Dealers Automated Quotations (NASDAQ) stock market lists approximately 3600 electric companies. In a stock market with large number of securities and trade volume, the computing time will be too long and even infeasible by using traditional quadratic programming method to obtain the optimal portfolio. In this article, the Cauchy–Schwarz Maximization (CSM) and the concept of Kuhn–Tucker (KT) conditions were applied to obtain a closed-form solution for the optimal Sharpe ratio portfolio. This approach is more efficient than traditional quadratic programming method via time and process savings.

This article is organized as follows. We first introduce the background and purpose of this study. Next we illustrate the model of optimal portfolio selection by maximizing the Sharpe ratio, followed by the discussion on a novel application of CSM and KT conditions for finding the closed-form solution of the model. A real world data will then be used to confirm the proposed closed-form solutions. Two scenarios including short sales allowed and short sales disallowed will be discussed. Moreover, we will use the proposed algorithms to find the optimal Sharpe ratio portfolio including 250 securities of Standard & Poor 500 (S&P 500), which is typically not possible by using traditional quadratic programming approaches. The last section provides conclusions and suggestions for future research.

II. Optimal Portfolio Selection by Maximizing the Sharpe Ratio

The Markowitz mean-variance model states that if the portfolio consists of $n$ securities, then its efficient
Optimal mean-variance portfolio selection using Cauchy–Schwarz maximization

The traditional methods to resolve the quadratic programming problems, such as Equations 1 and 3, usually have been related to Lagrange functions. When the number of securities (n) in portfolio is large, the resolving processes for Lagrange functions may become cumbersome. Moreover, the typical software used to solve linear or quadratic programming, such as LINDO or LINGO, is limited to 200 variables. A closed-form solution for Equation 3 is thus desirable.

III. Application of CSM and KT conditions

The extended Cauchy–Schwarz inequality in multivariate statistics gives rise to the following maximization lemma (Johnson and Wichern, 1992).

Lemma: Let \( B \) be a positive definite matrix and \( d \) be a given vector. Then for an arbitrary nonzero vector \( x \),

\[
\text{Max } \frac{x^T B x}{x^T d} = \frac{d^T B^{-1} d}{d^T d} \text{ with the maximum attained when } x = c B^{-1} d \quad \text{for any constant } c \neq 0.
\]

Replaced \( B, d \) and \( x \) by \( S \) (the covariance matrix of securities), \( e \) (the vector of excess return rates of securities) and \( w \) (the investment weights matrix) respectively, the above Lemma implies,

\[
\text{Max}(w^T e / \sqrt{w^T S w})^2 = e^T S^{-1} e \quad (4)
\]

with the maximum attained when \( w = c S^{-1} e \), for any constant \( c \neq 0 \).

Note that \( (w^T e / \sqrt{w^T S w})^2 \) is the square of the Sharpe ratio. Normally, the portfolio with the maximum Sharpe ratio has a positive expected return rate. Hence, Equation 4 above should yield the same optimal portfolio as Equation 3 if a suitable constant \( c \) is selected to ensure \( \sum w_i = 1 \). In other words, if the covariance matrix and the excess return rates vector of securities are known, the closed-form solution of optimal Sharpe ratio portfolio with short sales allowed can be obtained by the following algorithm.

CSM algorithm

Let \( \hat{\mathbf{w}}^T = \mathbf{S}^{-1} \mathbf{e} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_n) \) to obtain the primary solution of the investment weights matrix. The optimal solution \( \mathbf{w} \) can be obtained after normalization. That is \( \mathbf{w} = c \hat{\mathbf{w}} \), where \( c = (1 / \sum \hat{w}_i) \).

The optimal solution given in CSM algorithm does not guarantee all investment weights are positive. In a financial market which does not allow short sales, the requirement of all investment proportions of securities to be positive is needed. In this case, we
propose to employ the concept of KT conditions to obtain the feasible portfolio with the highest Sharpe ratio as follows.

**CSM–KT-algorithm**

**Step 1:** Compute  
\[ \hat{\mathbf{w}}^T = S^{-1} \mathbf{e} = (\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_n) = [\hat{w}_+, \hat{w}_-]^T \]  
and obtain the primary solution of the investment weights vector, where \( \hat{w}_+ \) is the vector including all positive weights, and \( \hat{w}_- \) represents the vector including all negative weights.

**Step 2:** If the investment weights \( \hat{w}_i \) of the vector \( \hat{w} \) are all positive, then proceed to Step 4. Else proceed to Step 3.

**Step 3:** Let all of the weights in \( \hat{w}_- = 0 \) and exclude these securities from the portfolio, and proceed to Step 1 by considering the securities in \( \hat{w}_+ \) only.

**Step 4:** Normalize the optimal solution to  
\[ \mathbf{w} = c\hat{\mathbf{w}}, \]  
where  
\[ c = (1 / \sum_{i=1}^{n} \hat{w}_i). \]

We next provide the theoretical justification on the CSM–KT-algorithm. The algorithm of quadratic programming problem, such as Equation 1 or Equation 3 with \( w_i \geq 0 \) are based on the concept of advanced calculus called KT conditions. Let \( \theta \) represent the Sharpe ratio and tentatively ignore the constraints \( \sum_{i=1}^{n} w_i = 1 \) and \( w_i \geq 0 \). A maximum value of \( \theta \) can be found by taking the derivative of \( \theta \) with respect to each \( w_i \) and setting it equal to zero. This maximum is indicated by point M in Fig. 1(a) for positive solution or Fig. 1(b) for negative solution. When \( w_i \) must be nonnegative and the optimal solution is negative as demonstrated in Fig. 1(b), the maximum feasible value of \( \theta \) occurs at point M’ rather than M. In other words, if the unconstrained maximum of \( \theta \) occurs at a value of negative \( w_i \), the feasible maximum value of \( \theta \) should occur at the point with \( w_i = 0 \) (see, e.g. Elton and Gruber, 1995; Taha, 1997). Therefore, the optimal weights obtained by the CSM-algorithm should be adjusted by letting the \( w_i = 0 \) and be excluded from the optimal portfolio when \( w_i \) is negative. In addition, the remaining securities with positive weights should be recalculated by the CSM-algorithm to obtain the optimal portfolio.

**IV. Empirical Analysis**

Two scenarios will first be discussed in this section: short sales allowed and short sales disallowed. In financial area, ‘short sales’ is defined as the investors sell stocks that they do not own. In the first scenario, the example concerns a mutual fund, which invests in six Dow Jones industrial index stocks and allows short sales. In the second scenario, the same example is used but short sales are not allowed. Furthermore, in the third scenario, the proposed algorithms will be used to find the optimal Sharpe ratio portfolio including 250 securities of S&P 500, where the traditional method is infeasible.

**Short sales are allowed**

A mutual fund manager is assumed to decide to invest in six stocks in the Dow Jones industrial index, which are AT&T (T), International Business Machines (IBM), Hewlett-Packard (HPQ), Coca-Cola (KO), Wal-Mart (WMT) and Home Depot (HD). Furthermore, short sales are allowed in this case. Annual returns over the period 1983 to 2006 are considered to estimate the mean return rates and correlation matrix of the six securities. Historical data collected from *Bloomberg system* in Merrill Lynch International Bank Limited was used to construct the
return series. Table 1 presents the means and SDs of these six stock returns. Table 2 presents the correlation matrix of the return rates of six securities. Note that the covariance matrix can be obtained easily by the correlation matrix and SDs. Although the empirical data applied annual returns in this study, quarterly or monthly returns can be used in the proposed models and obtain the same results. The process of using quarterly or monthly returns is the same as that of using annual returns, therefore employing the data with higher frequencies to extract optimal portfolios does not affect the outcome.

The data in Tables 1 and 2 are then used to solve the quadratic programming problem in Equation 3. A risk-free rate of 5% was assumed.

The optimal investment weights of six stocks, the mean return rate and the SD of the portfolio with the maximum Sharpe ratio are obtained and displayed in the first column of Table 3. As a comparison, we next apply the CSM-algorithm to the same data in Table 1. The inverse of the covariance matrix is multiplied by the excess return rates vector of securities and the optimal investment weights vector of six stocks (T, IBM, HPQ, KO, WMT, HD) = (−0.01197, 0.00574, −0.00274, 0.02191, 0.01033, 0.00506), is obtained. For this case, the optimal investment weight of T and HPQ are negative, implying that investors should short sell AT&T and HPQ stocks to achieve the highest Sharpe ratio. The inverse of the sum of these weights, 0.02833, is 35.2983. To make the sum of the investment weights equal to one, the primary weights were therefore multiplied by 35.2983 and the results in the second column of Table 3 were obtained. This solution confirms that the proposed algorithm yields the same result as the quadratic programming.

### Short sales are not allowed

When short sales are not allowed, investment weights cannot be negative. However, the optimal investment proportions of the example in the first subsection include some negative numbers. The optimal investment weights of AT&T (T) and HPQ are −42.24% and −9.66%. This solution is infeasible when short sales are not allowed. To solve the problem by quadratic programming method, the data in Tables 1 and 2 are also applied to Equation 3 with a constraint $w_i \geq 0$. The third column of Table 3 presents the optimal investment weights of six stocks, the mean return rate and the SD of the portfolio with the maximum Sharpe ratio obtained by solving quadratic programming.

The CSM–KT-algorithm first obtained (in Step 1) the optimal investment weights (T, IBM, HPQ, KO, WMT, HD) = (−0.01197, 0.00574, −0.00274, 0.02191, 0.01033, 0.00506). Since the optimal investment weights of T and HPQ are negative, this solution is infeasible. Thus we set $w_1$ and $w_3$ be zero and recalculate the optimal portfolio composed by

### Table 1. The means and SD of the six stocks returns

<table>
<thead>
<tr>
<th>T</th>
<th>IBM</th>
<th>HPQ</th>
<th>KO</th>
<th>WMT</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>9.12</td>
<td>10.87</td>
<td>12.83</td>
<td>18.82</td>
<td>28.96</td>
</tr>
<tr>
<td>SD (%)</td>
<td>29.29</td>
<td>32.52</td>
<td>32.78</td>
<td>25.55</td>
<td>39.43</td>
</tr>
</tbody>
</table>

*Source: Merrill Lynch International Bank Limited.*

### Table 2. The correlation matrix of the six stocks returns

<table>
<thead>
<tr>
<th>T</th>
<th>IBM</th>
<th>HPQ</th>
<th>KO</th>
<th>WMT</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1.0000</td>
<td>0.1484</td>
<td>0.3076</td>
<td>0.3685</td>
<td>0.4482</td>
</tr>
<tr>
<td>IBM</td>
<td>0.1484</td>
<td>1.0000</td>
<td>0.0382</td>
<td>−0.1451</td>
<td>0.2714</td>
</tr>
<tr>
<td>HPQ</td>
<td>0.3076</td>
<td>0.0382</td>
<td>1.0000</td>
<td>0.2990</td>
<td>0.3634</td>
</tr>
<tr>
<td>KO</td>
<td>0.3685</td>
<td>−0.1451</td>
<td>0.2990</td>
<td>1.0000</td>
<td>0.2756</td>
</tr>
<tr>
<td>WMT</td>
<td>0.4482</td>
<td>0.2714</td>
<td>0.3634</td>
<td>0.2756</td>
<td>1.0000</td>
</tr>
<tr>
<td>HD</td>
<td>0.3802</td>
<td>0.1025</td>
<td>0.4388</td>
<td>0.2796</td>
<td>0.7542</td>
</tr>
</tbody>
</table>

### Table 3. The maximum Sharpe ratio portfolio of the six stocks

<table>
<thead>
<tr>
<th>Short sales allowed</th>
<th>CSM-algorithm</th>
<th>Short sales not allowed</th>
<th>CSM–KT-algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>−0.4224</td>
<td>−0.4224</td>
<td>0</td>
</tr>
<tr>
<td>IBM</td>
<td>0.2027</td>
<td>0.2027</td>
<td>0.1355</td>
</tr>
<tr>
<td>HPQ</td>
<td>−0.0966</td>
<td>−0.0966</td>
<td>0.5153</td>
</tr>
<tr>
<td>KO</td>
<td>0.7733</td>
<td>0.7733</td>
<td>0.2292</td>
</tr>
<tr>
<td>WMT</td>
<td>0.3645</td>
<td>0.3645</td>
<td>0.1199</td>
</tr>
<tr>
<td>HD</td>
<td>0.1785</td>
<td>0.1785</td>
<td>22.30</td>
</tr>
<tr>
<td>$\mu$ (%)</td>
<td>28.90</td>
<td>28.90</td>
<td>22.67</td>
</tr>
<tr>
<td>$\sigma$ (%)</td>
<td>29.12</td>
<td>29.12</td>
<td>22.67</td>
</tr>
<tr>
<td>Sharpe ratio</td>
<td>0.822</td>
<td>0.822</td>
<td>0.763</td>
</tr>
</tbody>
</table>
other four stocks. In other words, AT&T and HPQ stocks are excluded from the portfolio. The updated weights are obtained (T, IBM, HPQ, KO, WMT, HD) = (0, 0.00456, 0, 0.01734, 0.00771, 0.00403). After normalization, the final result is displayed in the fourth column of Table 3. This solution confirms that the proposed algorithm yields the same result as the quadratic programming.

**A portfolio with large number of securities**

The proposed algorithms can obtain the optimal Sharpe ratio portfolio with a large number of securities. This is typically unsolvable by using traditional quadratic programming. For example, the number of variables is limited to be 200 (or less) by LINDO or LINGO software. Moreover, constructing the objective function of Equation 3 is troublesome. The variance of the portfolio will include 250(250 + 1)/2 terms for 250 stocks.

Since our proposed solution has a closed-form solution, we can easily construct this portfolio by considering daily returns of 250 stocks of S&P500 over the year of 2006. All data are obtained from Bloomberg system. By implementing the CSM-algorithm, we can obtain the weights of the optimal portfolio and this portfolio has \((\mu, \sigma) = (11.86\%, 0.21\%)\) with Sharpe ratio 56.48. Apart from some straightforward calculation, all we need is to solve the inverse of a 250 \( \times \) 250 covariance matrix. This can be done by many computing tools. MATLAB 7.0 software was used in this study.

**V. Conclusion**

Traditionally, financial analysts can revise the objective function of the Markowitz mean-variance portfolio model and resolve quadratic programming to obtain the maximum Sharpe ratio portfolio. This article proposed simple algorithms with closed-form solution by applying the CSM and the concept of KT conditions to decide optimal portfolios. The scenarios including short sales allowed and disallowed were discussed. A real world example demonstrated that the CSM-algorithm and the CSM–KT-algorithm correctly generate the closed-form solution for the portfolio with the highest Sharpe ratio. The proposed approach is more efficient than traditional quadratic programming method. Furthermore, in a stock market with large number of securities and trade volume such as NYSE or NASDAQ, the computing process for obtaining the optimal portfolio will be very complicated by using traditional quadratic programming methods. Therefore, the proposed CSM-algorithm and CSM–KT-algorithm can be important tools for seeking the optimal portfolio. This study used the CSM-algorithm to obtain the optimal portfolio with 250 securities. The calculation will be almost impossible via traditional quadratic programming approaches.

As in Kroll et al. (1984) and Campbell et al. (2001), the empirical examples of this article used historical data to estimate the means vector and covariance matrix of the return rates of securities. Forecasting technology, such as Vector Autoregressive (VAR) model, can be applied to obtain a more precise means vector and covariance matrix for returns of securities. Moreover, the CSM-algorithm and CSM–KT-algorithm provide closed-form solutions for the portfolio with maximum Sharpe ratio. The closed-form solutions are helpful to find the confidence intervals of the optimal investment weights of securities.

**Acknowledgements**

This research is supported by National Science Council of the Republic of China, Taiwan, under Contract No. NSC93 – 2416-H-110-009.

**References**


