2011 23rd IEEE International Conference on Tools with Artificial Intelligence

Portfolio Optimization through Data Conditioning and Aggregation

Elaine Wah
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA, USA
elaine.wah@ucla.edu

Yi Mei and Benjamin W. Wah
Department of Computer Science and Engineering
The Chinese University of Hong Kong
Hong Kong
{*ymei,wah*}*@cuhk.edu.hk*

***Abstract*—In this paper, we present a novel portfolio optimization method that aims to generalize the delta changes of future returns, based on historical delta changes of returns learned in a past window of time. Our method addresses two issues in portfolio optimization. First, we observe that daily returns of stock prices are very noisy and often non-stationary and dependent. In addition, they do not follow certain well-defined distribution functions, such as the Gaussian distribution. To address this issue, we first aggregate the return values over a multi-day period into an average return in order to reduce the noise of daily returns. We further propose a pre-selection scheme based on stationarity, normality and independence tests in order to select a subset of stocks that have promising statistical properties. Second, we have found that optimizing the average risk in a past window does not typically generalize to future returns with minimal risks. To this end, we develop a portfolio optimization method that uses the delta changes of aggregated returns in a past window to optimize the delta changes of future expected returns. Our experimental studies show that data conditioning and aggregation in our proposed method is an effective means of improving the generalizability while simultaneously minimizing the risk of the portfolio.**

***Keywords*-Portfolio selection; generalizability; data aggregation; stock pre-selection.**

I. Introduction

Given a set of available stocks, the goal of the portfolio selection problem in financial investment is to determine the optimal weight (*i.e.*, fraction of total investment) to assign to each stock in the portfolio, so that the future return is maximized and the risk is minimized.

For portfolio selection, Markowitz [1], [2] proposed a mean-variance model, in which the expected return and risk of the portfolio are formulated as the functions of the expected returns and covariance matrix of the stocks. Given the expected return and the covariance matrix, the portfolio is then optimized by a quadratic programming method.

In practice, the expected future return and the covariance matrix are unknown in advance and thus must be estimated based on historical data. In previous work, the sample mean and sample covariance matrix are often used to estimate the expected future return and covariance matrix. When the number of stocks is larger than the number of observations in the historical returns, the sample covariance matrix may be singular, and thus not applicable for quadratic programming. To address this issue, Ledoit and Wolf [3] improved the covariance-matrix estimation through shrinkage. Additionally, the estimation of the covariance matrix can be avoided by using the capital asset pricing model [4] or arbitrage pricing theory [5]. The effect of the estimation error is also considered in [6], [7], [8], [9].

In portfolio selection, there are two important issues that may lead to the poor performance of the future portfolio. First, the daily return has much noise and the portfolio has large variance of return when generalized to the next day. Moreover, the portfolio may not be generalizable due to poor statistical properties, such as non-stationarity, non-normality and dependence. Second, optimizing the average risk in a past window does not typically lead to future returns with similar risks. In this paper, we propose a two-step method involving data conditioning and aggregation to address these issues. The experimental studies on the S&P500 stocks and the entire US stock universe show that data conditioning and aggregation can improve the generalizability and reduce the risk of the portfolio.

The rest of the paper is organized as follows. Section II introduces the problems studied. We then present the method through data conditioning and aggregation in Section III, followed by the experimental results in Section IV. Finally, we give our conclusions and future work in Section V.

II. Problems Addressed and Approach

In this section, the problem statement that leads to data conditioning and aggregation is introduced.

A. Data Conditioning Through Pre-Selection

To facilitate the estimation of the expected risks of stock returns, it is often assumed that the distribution of the stock returns is multivariate normal and the parameters are fixed over time (i.e. stationary). Under this assumption, any weighted sum of the stock returns, e.g. the portfolio return, is normally distributed [10]. In addition, normal distributions have low kurtosis (i.e. skinny tails) and thus fewer extreme deviations from the mean. Stationarity, which implies that the mean and variance of a distribution do not change over time, is also important since the parameter estimation for a

1082-3409/11 $26.00 © 2011 IEEE
DOI 10.1109/ICTAI.2011.46

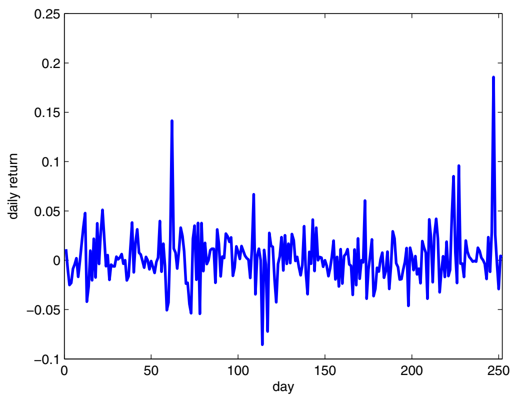


Figure 1. Daily returns of WDC between January 3, 2005, and December 30, 2005. The time series does not pass stationarity and normality tests.

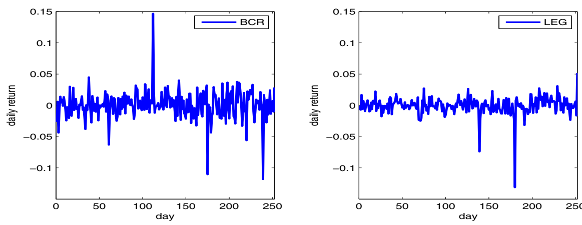


Figure 2. Daily returns of BCR and LEG between January 3, 2005, and December 30, 2005. The time series does not pass the independence test.

stationary time series over a long horizon is more precise than that of a non-stationary series [11].

Despite the promising properties of these assumptions, empirical studies have shown that daily stock returns are generally non-normal [12], [13] and non-stationary [11].

Figure 1 shows the daily returns of the Western Digital Corporation (WDC) in 2005, which is considered to be non-stationary and non-normal through statistical hypothesis testing. The time series has several large spikes that deviate significantly from the mean return over the time period. These extreme values cannot be predicted based solely on historical returns, and this type of behavior is more likely to appear for non-stationary and non-normal stocks, making it more difficult to accurately estimate the future returns for such stocks.

Dependence between stock returns is also important because the gains of diversification are mainly from variance reduction through combining independent or weakly correlated stock returns. Thus, if returns are dependent and have high variances, there is little gain from diversification in the portfolio. Figure 2 shows the return of two dependent stocks (BCR and LEG). Due to their dependence, they both exhibit a large negative return around day 175. A portfolio with non-zero and non-negligible allocations in these two stocks would exhibit a sharp decline in its generalized return if the learning window is before that day.

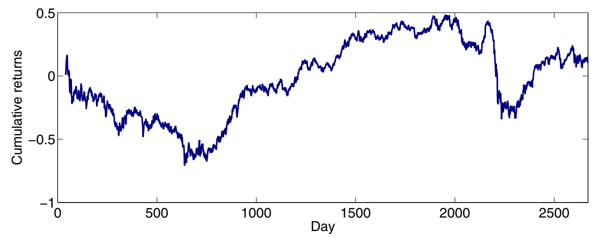


Figure 3. Cumulative generalized returns generated using the mean-variance approach with learning window size of 39 days on 24 stocks from the NASDAQ100 index between January 3, 2000 and October 8, 2010.

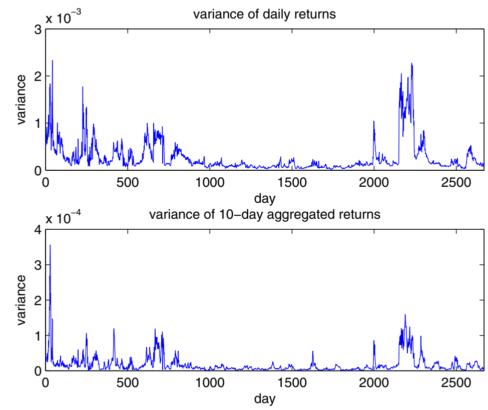


Figure 4. Variances of (a) daily returns and (b) 10-day moving average of daily returns in the learning window of the portfolio obtained by the mean-variance method with a window size of 39 days.

To address the issue that daily returns may be hard to generalize due to non-stationarity, non-normality and dependence, we propose in Section III-A to pre-select stocks that have well-defined and generalizable behavior, before constructing the portfolio. Specifically, we select from the entire set of stocks those that pass the statistical tests for stationarity, normality, and mutual independence.

B. Data Aggregation To Reduce Noise in Daily Returns

High variance in daily returns is another issue that leads to poor generalization. Based on a portfolio generated by applying the mean-variance method to 24 selected stocks from the NASDAQ100 index, Figure 3 illustrates that the portfolio does not generalize well to the next day. There are several extensive periods of consecutive negative returns, and the cumulative generalized return after nearly 10 years is barely higher than the initial investment.

Figure 4a shows the corresponding variances of the daily returns of the portfolio in the learning window. The variances are found to be high during periods with steep declines in the cumulative returns (e.g., around 2×10^{-3} at day 2200).

This example shows that the noise in daily returns is large, leading to large variances in the portfolio returns. However, returns aggregated over multiple days can have

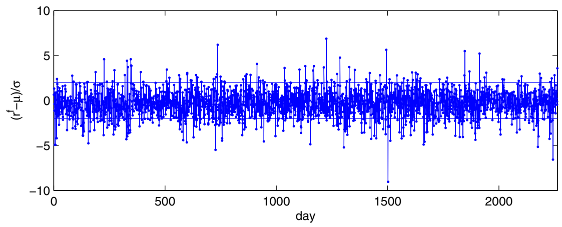


Figure 5. The term $(r^f - \mu)/\sigma$ of the portfolio obtained by the mean-variance method with window size of 44 days from January 2, 2001, to December 31, 2009.

smaller variances. A simplified justification of this behavior is as follows: when we assume that daily returns of different days are uncorrelated according to the random walk hypothesis [14]:

$$var\left(\frac{r_1 + \ldots + r_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} var(r_i) < \frac{\max_i\{var(r_i)\}}{n}.$$

Figure 4b shows the variances of the 10-day moving averages of daily returns; namely, $MA(t, 10) = \frac{1}{10}\sum_{i=0}^{9} r_{t-i}$, where r_t is the daily return at time t. It is an approximation of the true 10-day aggregated return $AG(t, 10) = \prod_{i=0}^{9}(1 + r_{t-i})^{1/10} - 1$, since r_t is so small and $r_t \approx \log(1 + r_t)$.

As the variances of the 10-day moving-average returns are much smaller (about 10% of the variances of daily returns), generalizing the portfolio to a multi-day period instead of a single day will reduce the risk of the portfolio. Based on this observation, we propose in Section III-B to aggregate the daily returns over multiple days.

C. Optimizing the Delta Changes of Daily Returns

When optimizing a portfolio, Markowitz' mean-variance method only minimizes the average risk of daily returns in a past window. However, the occurrence of extreme returns in the next day depends not only on the mean but also on the kurtosis of the distribution.

Figure 5 plots the normalized $z = (r^f - \mu_s)/\sigma_s$ of the portfolio found by the mean-variance method with window size of 44 days from Jan. 2, 2001, to Dec. 31, 2009, where r^f is the generalized daily return, and μ_s and σ_s are, respectively, the optimized mean and standard deviation of the portfolio in the past window. Under the normality assumption, the 95% confidence interval of the term is $[-2, 2]$. The figure shows that z is outside this interval in 621 out of 2,263 days and can even reach -10. This shows that the average risk in the past window cannot adequately capture the risk in the following day.

Based on the observation that the average risk in a past window does not generalize well to the risk of the return in the following day, we propose in Section III-B a portfolio optimization method that examines the delta changes of returns in a past window, with the objective of optimizing the delta changes of future returns.

III. Portfolio Optimization Based On Data Conditioning and Aggregation

We present in this section our portfolio optimization method using data conditioning and aggregation.

A. Stock Pre-selection

The goal of pre-selection is to increase the generalizability of those retained stocks. To this end, we eliminate those stocks that are significantly non-stationary, non-normal and dependent. Here, stationarity, normality and independence tests are employed to eliminate such stocks.

Stationarity. In general, stationarity of a time series implies that at least the mean and the variance are unchanged over time. To test the stationarity of the return time series with respect to its mean, we apply the Kwiatkowski, Phillips, Schmidt, Shin (KPSS) test [15], which tests the null hypothesis that an observable time series is stationary around a deterministic trend. Then, the variance-ratio test [16] is used to test for stationarity with respect to variance.

Normality. We apply the Lilliefors test [17], [18] to test the normality of each stock return time series and reject those that are not normally distributed. Such stocks should be eliminated because they have fat tails and high risk levels.

Independence. Since zero correlation between two random variables does not imply independence whereas two random vectors are independent iff their distance covariance is zero [19], we use the distance covariance as a measure of dependence and develop the following independence test for the pre-selection process.

Applying the independence test to the return time series results represented in an $N \times N$ binary matrix $\mathbf{V}$, where N is the number of stocks. An entry of $v_{i,j}$ of 1 indicates that the returns of stock i and j are significantly dependent, while 0 indicates that the null hypothesis cannot be rejected. Then, the largest subset of mutually "independent" stocks are selected from $\mathbf{V}$. This is equivalent to a maximum-clique problem in which stocks are "connected" if they pass the independence test, which is *NP*-hard [20]. Therefore, we propose a greedy iterative heuristic to select the stocks. In each iteration, we eliminate the stock that is dependent with the greatest number of other stocks by removing the corresponding column and row of $\mathbf{V}$. We repeat the process until all entries in the reduced $\mathbf{V}$ are zero.

As an example, Figure 6 shows the returns from the S&P500 set that pass all the statistical tests. In this case, the daily returns for those 27 retained stocks exhibit few extreme, high-magnitude changes in value, and they do not appear to follow any overall trend.

B. Aggregation Algorithm for Portfolio Optimization

Our proposed aggregation method estimates the average expected return over the next g days. While the mean-variance formulation assumes that the objectives of the investor are maximizing the generalized return on investment

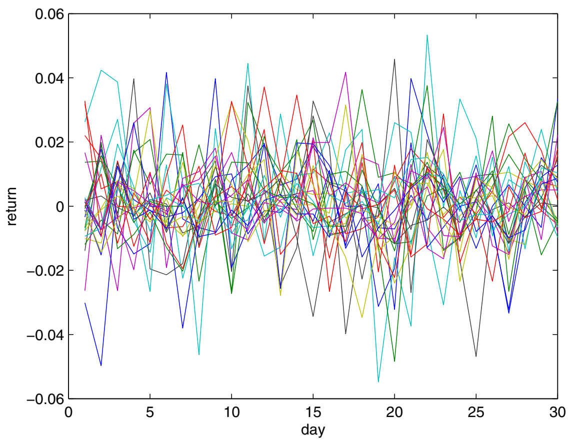


Figure 6. Daily returns of 27 stocks (out of 464 total) between May 25 and July 8, 2005, that pass stationarity, normality, and independence tests.

and minimizing the risk, the objective of our approach is to solely maximize mean return over the entire learning window. Risk is implicitly minimized by including a constraint to bound the maximum difference (or delta) between adjacent sample average expected returns.

Let $r_{i,j}$ be the daily return of stock i at day j, and R_j be the portfolio return at day j. The g-day sample averaged returns at day j are:

$$\overline{r_{i,j}} = \frac{1}{g}\sum_{k=0}^{g-1} r_{i,j-k} \quad \text{and} \quad \overline{R_j} = \frac{1}{g}\sum_{k=0}^{g-1} R_{j-k}.$$

At day t, comparisons are made for c pairs of adjacent non-overlapping averaged portfolio returns $\overline{R_{t-d-g}}$ and $\overline{R_{t-d}}$, $d = 0,\ldots,c-1$. The delta between each pair of the averaged returns is constrained to be no larger than the threshold δ. Using a corresponding learning window size of $2g+c-1$, the returns in the past $2g+c-1$ days are used for optimization (from day $t-2g-c+2$ to day t). The portfolio optimization model can be formulated as follows:

$$\max_{\boldsymbol{x}} \ \boldsymbol{f}^T\boldsymbol{x} \tag{1}$$

$$s.t.: -\delta\mathbf{1} \leq \boldsymbol{D}\boldsymbol{x} \leq \delta\mathbf{1} \tag{2}$$

$$\sum_{i=1}^{N} x_i = 1 \tag{3}$$

$$x_i \geq 0, i = 1,\ldots,N. \tag{4}$$

where N is the number of available stocks, and $\boldsymbol{x} = (x_1,\ldots,x_N)$ is the weight vector of the stocks. Here, $\boldsymbol{f} = (f_1,\ldots,f_N)$ is the vector of the average returns in the learning window:

$$f_i = \frac{1}{2g+c-1}\sum_{k=0}^{2g+c-2} r_{i,t-k}, \quad i = 1,\ldots,N,$$

where $\mathbf{1}$ is the $c\times 1$ vector of ones; δ is a pre-defined parameter indicating the bound of the difference between adjacent aggregated returns; and $\boldsymbol{D}$ is a $c\times N$ matrix of the difference between the c pairs of adjacent averaged returns of the N stocks. It is described as follows:

$$\boldsymbol{D} = \begin{pmatrix} \overline{r_{1,t}} - \overline{r_{1,t-g}} & \cdots & \overline{r_{N,t}} - \overline{r_{N,t-g}} \\ \overline{r_{1,t-1}} - \overline{r_{1,t-1-g}} & \cdots & \overline{r_{N,t-1}} - \overline{r_{N,t-1-g}} \\ \cdots & \cdots & \cdots \\ \overline{r_{1,t-c+1}} - \overline{r_{1,t-c+1-g}} & \cdots & \overline{r_{N,t-c+1}} - \overline{r_{N,t-c+1-g}} \end{pmatrix}$$

In our formulation, the objective (1) is to maximize the average portfolio return in the learning window. Constraint (2) indicates that the difference between each pair of adjacent averaged returns in the learning window cannot exceed the bound δ, and Constraints (3) and (4) are the basic constraints of portfolio selection.

To prevent negative portfolio return in the case that all the selected stocks have negative expected returns, a "cash" stock with a risk-free daily return r_f is included in the stock pool. In the worst case, all the weights are allocated to this risk-free stock.

After the portfolio at day t has been obtained, we set the generalized period from day $t+1$ to day $t+g$; that is, the portfolio is held for g days before the investment return is actualized. Consequently, the investor will not see the actualized return of the portfolio obtained at day t until day $t+g$. The generalized portfolio return is calculated as:

$$\overline{R_t^g} = \frac{1}{g}\sum_{j=1}^{g} R_{t+j}.$$

Finally, to increase the diversity of the portfolio and to prevent the situation in which one or a few stocks contribute significantly to the portfolio than the remaining stocks, we impose constraints on the maximum weight of each stock. We consider the following two types of constraints.

a) *Equal-weight constraints.* In these constraints, the maximum weight of each stock is set to an equal value depending on the number of stocks with past average return larger than the risk-free return r_f:

$$x_i \leq 1/K, \quad i = 1,\ldots,N, \tag{5}$$

where $K = ||\{i \in \{1,\ldots,N\} | f_i > r_f\}||$ is the number of stocks with past average return larger than r_f.

b) *Equal-contribution constraints.* In these constraints, the maximum weight of each stock is set in such a way that its maximum contribution to the portfolio is equal, where the contribution is represented by the weighted standard deviation:

$$x_i \leq \begin{cases} 1 & \text{if } f_i \leq r_f \\ X_i & \text{if } f_i > r_f. \end{cases} \tag{6}$$

Here, X_i satisfies the following conditions:

$$X_i s_i = X_j s_j, \quad \forall f_i > r_f, \quad f_j > r_f, \tag{7}$$

$$\sum_{f_i > r_f} X_i = 1, \tag{8}$$

where s_i is the standard deviation of stock i, and$X_i s_i$ is the maximum contribution of stock i to the portfolio. By solving (7) and (8), the closed-form solution of X_i is:

$$X_i = \frac{\frac{1}{s_i} \prod_{f_i > r_f} s_i}{\sum_{f_i > r_f} \frac{1}{s_i} \prod_{f_i > r_f} s_i}, \quad \forall f_i > r_f. \tag{9}$$

Figure 7 illustrates the effect of the equal-weight and equal-contribution constraints. The figures show that, with no constraint on the maximum weight on the stocks, there are generally two stocks in the portfolio. In contrast, the portfolios have much better diversification with the equal-weight or the equal-contribution constraints.

IV. Experimental Studies

In this section, we evaluate the efficacy of data conditioning and aggregation in our proposed method and investigate their limitations. In our experiments, we use the S&P500 data set and the stocks with market capitalization no less than US$500 million in the entire US market from January 2, 2001, to December 31, 2009 (totally 2263 trading days). During this period, there are about 450 S&P500 stocks and 2000 eligible stocks in the entire US stock set. Figure 8 shows the number of retained stocks in S&P500 set and the entire US market from January 2, 2001, to December 31, 2009. It is seen that the number of retained stocks is small, especially for the S&P500 stocks. To increase the flexibility of portfolio optimization, the entire US stock set is more desirable than the S&P500 stock set.

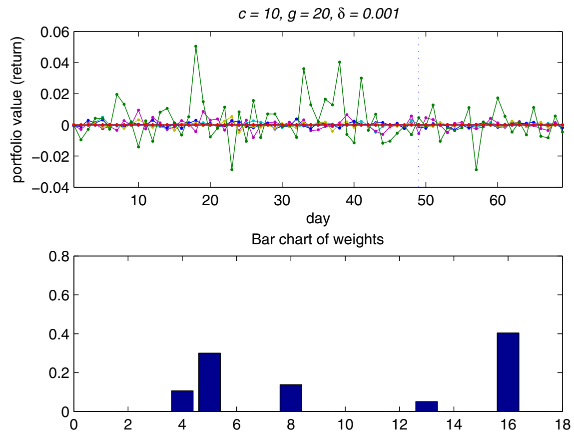
A. Effects of Stock Pre-selection

For increased selectivity, we set the confidence levels of the KPSS test and the variance-ratio test to 0.1, and those of the Lilliefors test and independence test to 0.15. We investigate the effects of stock pre-selection by comparing the portfolios obtained from the stocks with and without pre-selection. In our portfolio optimization, we set the parameters as follows: $c = 10$, $g = 20$, $\delta = 0.001$ and $r_f = 4 \times 10^{-6}$. As a result, the learning window has $c + 2g - 1 = 49$ days.

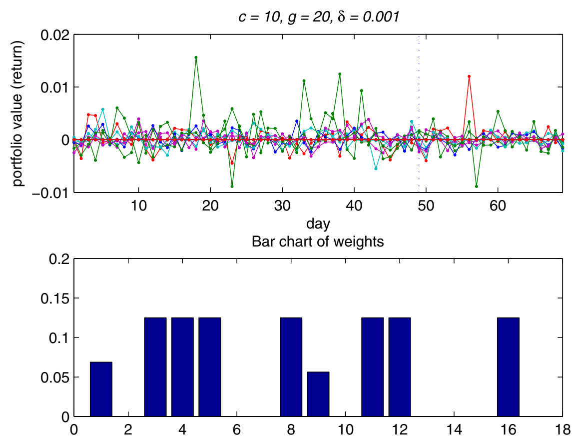
Figure 9 shows the cumulative returns of the portfolio with and without pre-selection in the S&P500 and the entire stock sets. It can be seen that for both cases, including the pre-selection before optimization leads to a much smoother cumulative-return curve. For example, during the period around the day 1850, although the cumulative returns of both methods have negative slopes, the magnitude of those of the method with pre-selection is much smaller.

B. Effects of Aggregation

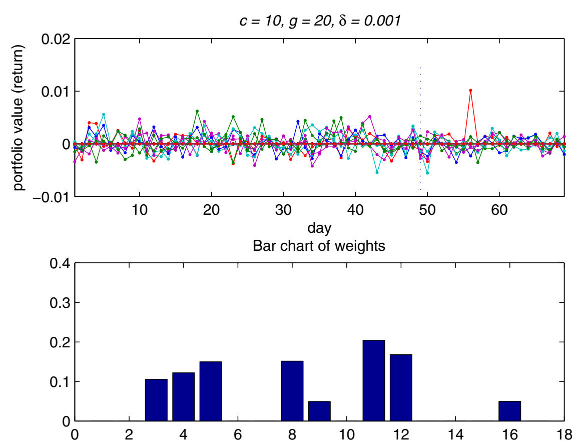
To investigate the effect of aggregating the returns on the generalized performance of the portfolio, Figure 10 compares the generalized aggregated returns of our proposed



(a) The portfolio returns with no constraint



(b) The portfolio returns with the equal-weight constraints



(c) The portfolio returns with the equal-contribution constraints

Figure 7. Applying the delta method with $c = 10$, $g = 20$ and $\delta = 0.001$ on those pre-selected S&P500 stocks with the learning period from Jan. 3, 2005, to March 14, 2005, and the generalization period from March 15, 2005, to April 12, 2005.

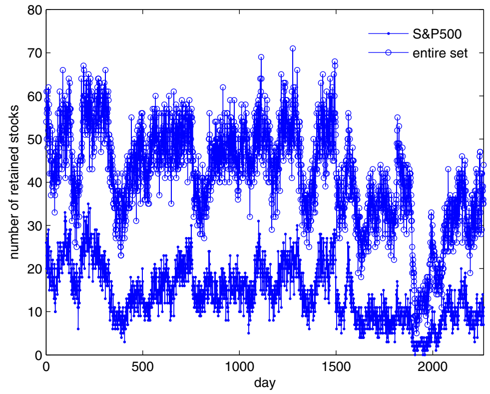


Figure 8. The number of retained stocks in the S&P500 set and the entire US market from January 2, 2001, to December 31, 2009.

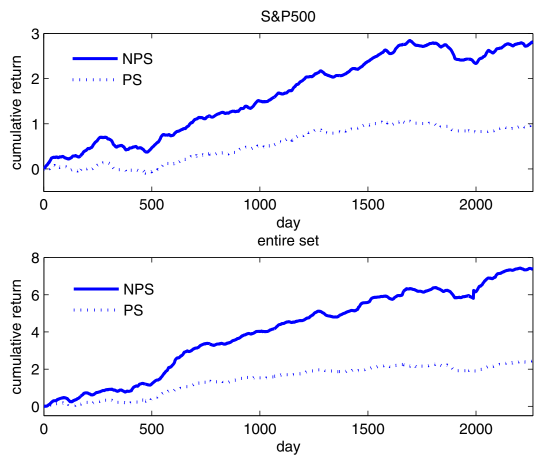


Figure 9. Cumulative returns of the portfolios (a) with and (b) without pre-selection in the S&P500 and the entire stock sets between January 2, 2001, and December 31, 2009.

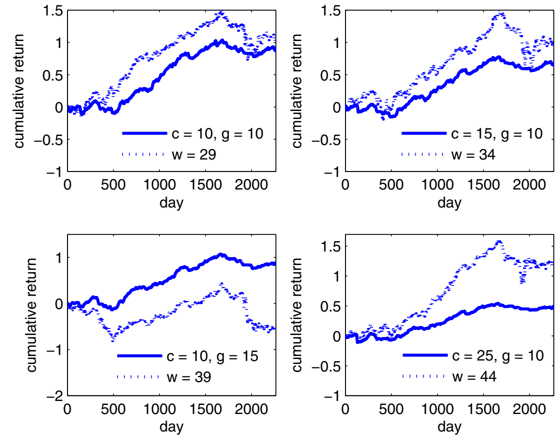


Figure 10. Comparison between the cumulative aggregated returns of our proposed method and the generalized daily returns of the mean-variance method with a corresponding window size of $2g + c - 1$ on the S&P500 data from January 2, 2001, to December 31, 2009, using four different combinations of parameters.

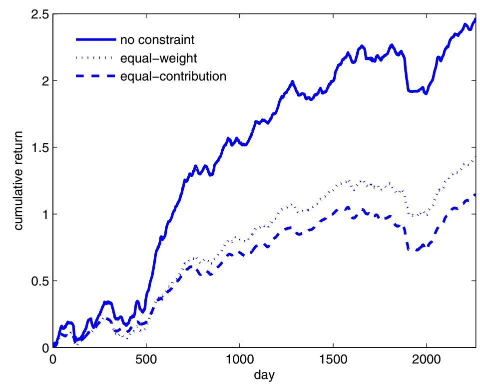


Figure 11. Cumulative returns of the three compared methods on the entire stock set from January 2, 2001, to December 31, 2009.

method and the generalized daily returns of the mean-variance method on the S&P500 data from January 2, 2001, to December 31, 2009. The graphs show that the cumulative-return curve with data aggregation is much smoother with smaller risks. These verify the efficacy of data aggregation in reducing the portfolio risk.

C. Effects of Constraining the Maximum Weights of Stocks

In this section, we compare the three ways of setting weights: the one with no constraint on the maximum weight of stocks, the one with equal-weight constraints, and the one with equal-contribution constraints on all the stocks. Figure 11 shows the cumulative returns of the three methods compared in the period between January 2, 2001, and December 31, 2009. For the methods with constraints on equal-weight and equal-contribution, the cumulative return curves are smoother than that of the method with no constraint on weights. The equal-weight method is also better than the equal-contribution method, as it has similar variance but higher final cumulative return.

D. Effects of Constraining Delta Changes

Figures 12 and 13 show, respectively, the cumulative returns and the deltas generalized to the next period of our proposed method with $\delta = 0.001$, 0.0005 and 0.0001 on the entire stock set from January 2, 2001, to December 31, 2009. With a smaller δ, there is more diversification in the portfolio, leading to a smoother cumulative return curve and a smaller generalized delta, but at a cost of smaller cumulative returns. However, the deltas generalized to the future still deviate significantly from the prescribed δ's in the constraints (and predominantly negative). This is attributed to over-constraining the stocks in the past window in order to obtain high returns and the tendency of the high returns of these over-constrained stocks to drop in the future.

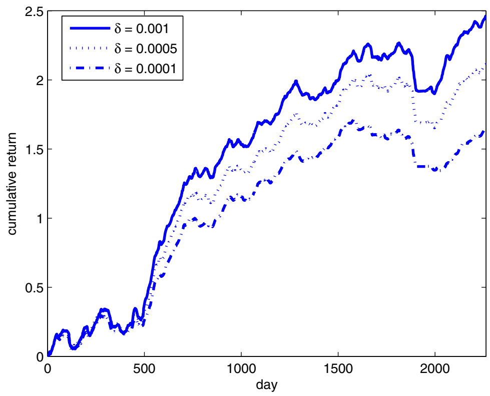


Figure 12. Cumulative returns of the methods with different δ's on the entire stock set from January 2, 2001, to December 31, 2009.

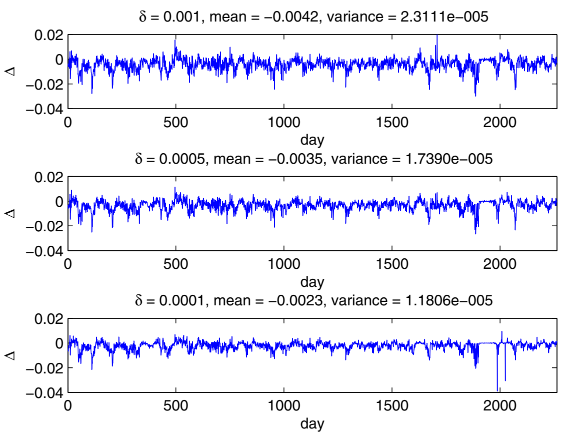


Figure 13. Deltas generalized to the future aggregation period of our proposed method with (a) $\delta = 0.001$, (b) $\delta = 0.0005$ and (c) $\delta = 0.0001$ on the entire stock set from January 2, 2001, to December 31, 2009.

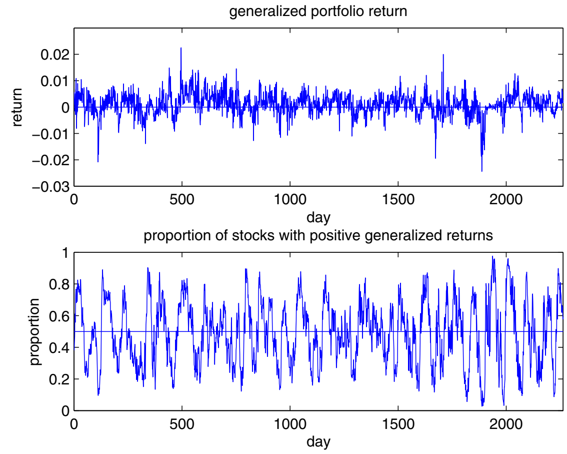


Figure 14. (a) The generalized portfolio returns and (b) the proportion of stocks with positive generalized returns from January 2, 2001, to December 31, 2009.

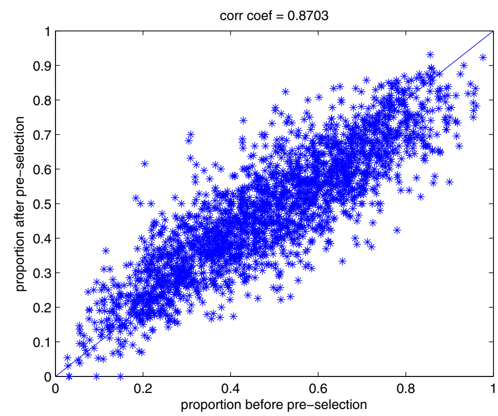


Figure 15. Proportions of stocks with positive generalized returns before and after pre-selection from January 2, 2001, to December 31, 2009.

E. Limitations of Our Proposed Approach

We have demonstrated that the pre-selection and aggregation method can improve the generalization and reduce the risk of the portfolio. However, the returns of our proposed methods are not always positive. In some periods (such as around days 100, 300, 1300 and 1900 in Figure 11), the cumulative-return curves of all the methods fall. That is, the portfolios obtained by all the methods have negative generalized aggregated returns during these periods.

One reason for the above behavior is that none of the methods take into account the general market conditions in its optimization. It is found that in the aforementioned days, most of the stocks have negative generalized aggregated returns. Figure 14 shows the generalized returns obtained by the aggregation method with no constraint on the maximum weights of the stocks and the proportion of stocks with positive generalized aggregated returns in the entire stock set from January 2, 2001, to December 31, 2009. It can be seen that when the generalized portfolio return is negative, the corresponding proportion of stocks with positive generalized return is generally smaller than 0.5, and vice versa. Since the pre-selection scheme does not consider the proportion of those retained stocks with positive generalized returns, the proportion would not change much after the pre-selection.

Figure 15 shows the relationship between the proportions of stocks with positive generalized aggregated returns before and after pre-selection. It shows that the two proportions are highly correlated (with correlation coefficient of 0.8703). Since the points are located around the line $y = x$ with no trend, pre-selection does not lead to a higher proportion of stocks with positive generalized returns. It can be easily understood that, when most of the available stocks have positive (*resp.* negative) generalized returns, the obtained portfolio will very likely have positive (*resp.* negative) generalized return regardless of the optimization method. To further improve the proportion of stocks with positive generalized returns, it is important to develop a pre-selection method that examines the market condition as well.

V. Conclusions and Future Work

In this paper, we have addressed the generalization of portfolio optimization and the high variances of daily returns by proposing a portfolio-selection method based on data conditioning and aggregation. Based on pre-selection, we select a subset of stocks that are more generalizable from the entire stock universe and apply a new portfolio optimization method to generalize the portfolio return to a multi-day period in the future. The experimental results demonstrate the efficacy of data conditioning and aggregation in improving the generalization and in reducing the risk of the portfolio.

In the future, we plan to study the following issues.

- Without examining the market condition, pre-selection can minimally improve the generalized returns of the selected stocks. To this end, we plan to study those factors and market conditions that may affect such proportion for inclusion in pre-selection and to determine in what cases pre-selection is effective.
- After pre-selection, the number of retained stocks is relatively small (only around 50 retained from a universe of 2000 stocks). To increase the flexibility of stocks used in portfolio optimization, we plan to relax pre-selection in order to increase the pool of retained stocks.

References

[1] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[2] ——, *Portfolio selection: efficient diversification of investments*. Wiley, 1991.

[3] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, 2003.

[4] W. Sharpe, "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, vol. 19, no. 3, pp. 425–442, 1964.

[5] S. Ross, "Arbitrage theory of capital asset pricing," *Journal of Economic Theory*, vol. 13, no. 3, pp. 341–360, 1976.

[6] R. Klein and V. Bawa, "The effect of estimation risk on optimal portfolio choice," *Journal of Financial Economics*, vol. 3, no. 3, pp. 215–231, 1976.

[7] P. Jorion, "Bayes-stein estimation for portfolio analysis," *Journal of Financial and Quantitative Analysis*, vol. 21, no. 3, pp. 279–292, 1986.

[8] R. Michaud, "The Markowitz optimization enigma: is 'optimized' optimal?" *Financial Analysts Journal*, vol. 45, no. 1, pp. 31–42, 1989.

[9] R. Kan and G. Zhou, "Optimal portfolio choice with parameter uncertainty," *Journal of Financial and Quantitative Analysis*, vol. 42, no. 3, pp. 621–656, 2007.

[10] S. Kon, "Models of stock returns – a comparison," *Journal of Finance*, vol. 39, no. 1, pp. 147–165, 1984.

[11] E. Elton, M. Gruber, S. Brown, and W. Goetzmann, *Modern portfolio theory and investment analysis*. Wiley-India, 2009.

[12] E. Fama, *Foundations of finance*. Basic Books New York, 1976.

[13] S. J. Brown and J. B. Warner, "Using daily stock returns: The case of event studies," *Journal of Financial Economics*, vol. 14, no. 1, pp. 3–31, 1985.

[14] E. Fama, "Random walks in stock market prices," *Financial Analysts Journal*, vol. 21, no. 5, pp. 55–59, 1965.

[15] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root:: How sure are we that economic time series have a unit root?" *Journal of Econometrics*, vol. 54, no. 1-3, pp. 159–178, 1992.

[16] A. Lo and A. MacKinlay, "The size and power of the variance ratio test in finite samples: A Monte Carlo investigation," *Journal of Econometrics*, vol. 40, no. 2, pp. 203–238, 1989.

[17] H. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.

[18] ——, "On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown," *Journal of the American Statistical Association*, vol. 64, no. 325, pp. 387–389, 1969.

[19] G. Székely, M. Rizzo, and N. Bakirov, "Measuring and testing dependence by correlation of distances," *Annals of Statistics*, vol. 35, no. 6, pp. 2769–2794, 2007.

[20] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to algorithms*. The MIT press, 2001.