RESEARCH SUPPORT
UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

SEPTEMBER 1997

OPTIMAL SELECTION OF HEDGE AND INDEXED PORTFOLIOS

WORKING PAPER #9712-19

BY

MARTIN R. YOUNG
UNIVERSITY OF MICHIGANBUSINESS SCHOOL

JOHN CHENG
UNIVERSITY OF MICHIGAN

Optimal Selection of Hedge and Indexed Portfolios

Martin R. Young* John Cheng†

September, 1997

*Department of Statistics & Management Science, University of Michigan Business School, Ann Arbor, MI 48109. Email: myoung@umich.edu, Phone: 313-936-1332, Fax: 313-936-0274.

†Department of Economics, University of Michigan, Ann Arbor, MI 48109. Email: jcheng@umich.edu

Optimal Selection of Hedge and Indexed Portfolios

Abstract

A common objective in financial management is the construction of a portfolio whose returns closely track the returns of some other target asset; hedging, and indexed investing are two such applications. It is often necessary that the constructed portfolio be composed of a small number of underlying assets, in order to limit transaction and management costs. In this paper, a mixed integer programming algorithm is presented for choosing a subset of assets, and the corresponding allocations for those assets, such that the portfolio thus composed optimally tracks a given target. The algorithm is designed to be robust with respect to outliers, a critical feature given the familiar kurtosis of financial time series data. The algorithm is evaluated on historical stock return data, and it is seen that accurate out-of-sample tracking of a target can potentially be achieved with a small number of optimally selected assets.

I Introduction

A common financial engineering objective is the creation of a synthetic portfolio whose returns closely match those of some other target asset. Hedging is one possible motivation for wishing to replicate a target returns series; Anderson and Danthine (1981) and Herbst and Marshall (1995) discuss asset hedging with multiple instruments. Indexed investing offers another setting in which one attempts to create a portfolio whose returns are nearly identical to those of some other asset or portfolio.

The costs of creating and managing an indexing or hedging portfolio are minimized partly by limiting the number of assets purchased to compose the portfolio. Choosing the optimal subset of assets to include in the portfolio can be a challenging task; the choice depends not just on the marginal correlations between the hedging assets and the target asset, but rather on the *partial* correlations. This implies that the correlations between the hedging assets need to be considered, just as is the case in traditional portfolio optimization (Markowitz 1952). If the number of candidate assets for forming the hedge or index portfolio is sizable, then the number of pairwise correlations becomes larger than can be sensibly processed, other than through automated methods. The purpose of this article is to describe a mixed integer programming algorithm for achieving the dual objectives of portfolio replication: namely, tracking the target as closely as possible, while limiting the number of assets included in the portfolio.

The algorithm introduced in this paper uses an objective criterion specially suited to the frequently heavy-tailed, or non-Gaussian, character of asset returns data (Fama 1965; McCulloch 1978; Kon 1984). Thus, the portfolio allocation is robust with respect to the large outliers that can appear in financial time series data.

Section II describes the algorithm for optimal indexed portfolio selection. Section III describes applications of this technique to historical data on international stock returns, and to data on returns of firms trading on the New York Stock Exchange; here, it is seen that the robustness built into the algorithm does lead to performance gains, and that, indeed, a small number of optimally selected assets can match a target portfolio more accurately, out-of-sample, than does a portfolio composed of a larger set of assets, but with weights chosen without optimization. Section IV gives further discussion, including references to related work, and directions for future research.

II The Algorithm

Let m_t denote the return on a target portfolio in time period t, and y_{jt} be the return on individual asset j at time period t, $j = 1, \dots, N$. Suppose that observations of m_t and y_{jt} are available for $t = 1, \dots, T$; these observations will usually be historical returns, but could also be realizations from a Monte Carlo model which simulates future returns. The objective of index optimization is to choose a subset of assets S, $S \subseteq \{1, \dots, N\}$, and associated portfolio allocations w_j, $j \in S$ for these assets, so that the portfolio returns $p_t = \sum_{j \in S} w_j y_{jt}$ "closely match" the returns of the target portfolio m_t, over the observation period $t = 1, \dots, T$.

Least Absolute Deviations Estimation

The choice of the particular metric for determining the "closeness" of the match between a candidate index and the target portfolio over the observation period is in fact significant, due to the leptokurtic, or heavy-tailed, nature of financial return data (e.g., Kon 1984). The commonly used least squares (LS) procedure is based on the L_2 metric, which may be written as $L_2 = \sum_{t=1}^{T} (m_t - p_t)^2$; since the residuals in the L_2 norm are squared, the LS procedure can be very substantially affected by a single observation in which the target portfolio m_t and the index portfolio p_t deviate. Thus, a selection procedure which seeks to minimize an L_2 norm may be very sensitive to the outliers commonly observed in financial data. Instead, the algorithm to be presented in this paper will use the "least-absolute-deviations" (LAD) criterion, based on the L_1 norm: $L_1 = \sum_{t=1}^{T} |m_t - p_t|$. The LAD criterion is frequently used in statistical modeling, to achieve robustness against gross outliers. The median, for example, is an LAD estimator of the location of a distribution (Lehmann 1983). As another example of LAD modeling, Figure 1 displays a scatterplot of the monthly returns of a value weighted stock market portfolio ("VWI" - value weighted index) vs. the corresponding monthly stock returns for the Exxon Corporation, during 1991–1992, as well as the LS and LAD best-fitting lines for the scatterplot. The LAD method chooses the intercept β_0 and slope β_1 to minimize $\sum_{i=1}^{24} |y_i - \beta_0 - \beta_1 x_i|$, while the LS method chooses these parameters to minimize $\sum_{i=1}^{24} (y_i - \beta_0 - \beta_1 x_i)^2$. The large outlier appears to significantly influence the LS estimate; the LS estimate of the slope of the regression line is 0.42, while the LAD estimate is 0.22. Dielman and Pfaffenberger (1982) and Gonin and Money (1989) describe the advantages of the LAD metric over the LS metric in terms of robustness to outliers, in the context of linear regression modeling. Hull (1997) discusses the relationship between regression slope coefficients and optimal hedge ratios.

[Insert Figure 1 Here]

Linear–Integer Programming Formulation

Consider now the problem of choosing portfolio allocations w_j for the N securities such that the constructed portfolio closely matches the specified target portfolio m_t, in the L_1 norm, but with the additional restriction that at most $M \leq N$ securities get non-zero allocation. It will now be shown that this problem can be expressed as a mixed linear integer program.

Define ϵ_t^+ and ϵ_t^- as the positive and negative deviations between the target portfolio m_t and the indexed portfolio p_t: $m_t = p_t + \epsilon_t^+ - \epsilon_t^-$, with $\epsilon_t^+ = \max(m_t - p_t, 0)$ and $\epsilon_t^- = \max(p_t - m_t, 0)$. Define the indicator variables z_j as

$$z_j = \begin{cases} 1 & \text{if } w_j > 0 \\ 0 & \text{if } w_j = 0. \end{cases} \tag{1}$$

Then the optimization problem of choosing the non-negative portfolio weights w_j such that the portfolio $p_t = \sum_{j=1}^{N} w_j y_{jt}$ is as close as possible to the target m_t, and such that at most M of the w_j are non-zero, can be written as:

$$\min_{\mathbf{w},\mathbf{z},\epsilon^+,\epsilon^-} \sum_{t=1}^{T} \epsilon_t^+ + \epsilon_t^- \tag{2a}$$

subject to

$$p_t = \sum_j w_j y_{jt}, \quad t = 1, \ldots, T \tag{2b}$$

$$m_t - p_t = \epsilon_t^+ - \epsilon_t^-, \quad t = 1, \ldots, T \tag{2c}$$

$$\sum_{j=1}^{N} w_j = 1 \tag{2d}$$

$$0 \leq w_j \leq 1, \quad j = 1, \ldots, N \tag{2e}$$

$$w_j \leq z_j, \quad j = 1, \ldots, N \tag{2f}$$

$$0 \leq z_j \leq 1, \quad z_j \text{ integer}, \quad j = 1, \ldots, N \tag{2g}$$

$$\sum_{j=1}^{N} z_j \leq M. \tag{2h}$$

The objective is the sum of the absolute deviations between p_t and m_t. Equations (2f) and (2g) guarantee that the indicator variables z_j will obey definition (1); equation (2h) guarantees that at most M of the securities will have $z_j = 1$, and hence will have non-zero weight.

Equation (2e) forces allocations to be positive, and thus imposes a restriction against short selling. In order to permit a short position in, say, asset k, an $(N+1)$st asset can be added to the set of candidate assets, with returns defined by $y_{(N+1)t} = -y_{kt}$, $t = 1, \ldots, T$.

Equations (2a)–(2h) represent a mixed integer program, which can be solved using branch–and–bound or branch–and–cut methods (Lucena and Beasley 1996). There is special structure to the program which can be exploited to improve efficiency, since it is known that, at the optimal solution, either ϵ_t^+ or ϵ_t^- must be zero. This complementary slackness condition suggests that if ϵ_t^+ is entered into a simplex basis, then ϵ_t^- may be removed, since the variables cannot both be basic at the optimum (Barrodale and Roberts 1978; Armstrong and Godfrey 1979). In addition, specialized integer programming algorithms can be used for treating the fixed–charge type variables z_j; see, e.g., McKeown (1981) and Suhl (1985).

Extensions

Equations (2a)–(2h) define the fundamental algorithm for identifying the portfolio weights w_j which provide optimal matching between the portfolio p_t and the target portfolio m_t. There are some modifications to this formulation which may be useful in certain applications. Kariya (1993) suggests adding a constraint on the mean return of the matching portfolio; e.g.,

$$\sum_{j=1}^{N} w_j \bar{y}_j \geq G, \tag{2i}$$

for some value G, where $\bar{y}_j = T^{-1} \sum_{t=1}^{T} y_{jt}$. The associated integer program would then attempt to achieve a tradeoff between high average portfolio return, and low average deviation from the target, a tradeoff similar to the one seen in traditional mean-variance portfolio optimization (Markowitz 1952), as well as recent variations (Yamakazi and Konno 1991), and related to the "cost of the hedge" concept in Herbst and Marshall (1995).

A similar extension to program (2) can be obtained by generalizing the objective as:

$$\min_{\mathbf{w},\mathbf{z},\epsilon^+,\epsilon^-} \sum_{t=1}^{T} \tau\epsilon_t^+ + (1-\tau)\epsilon_t^-, \tag{2a'}$$

where τ is a real number in $(0,1)$; for $\tau \neq 0.5$, this has the effect of differentially penalizing positive and negative deviations between the target and matching portfolios. This optimization problem can be interpreted as a *quantile regression*: the portfolio p_t is selected to match the τ'th percentile of m_t conditional on $y_{1t}, \ldots, y_{Nt}$ (Koenker and Portnoy 1997). For $\tau = 0.5$, the matching portfolio is selected as the *conditional median* of m_t given $y_{1t}, \ldots, y_{Nt}$.

Given that M assets are to be included in the matching portfolio p_t, one may wish to ensure that wealth is distributed relatively evenly over the M assets; i.e., that the matching portfolio is suitably "diversified". This consideration can be incorporated by, for example, adding to the portfolio optimization problem (2) the following constraints:

$$w_j \leq 1 - \alpha \left(\frac{M-1}{M} \right), \quad j = 1, \ldots, N \tag{2j}$$

for some α in $[0, 1]$. Choosing $\alpha = 0$ in (2j) is effectively non–binding, since it is redundant with (2e), while choosing $\alpha = 1$ forces each of the M securities included in the portfolio to have equal weight, $1/M$, and thus forces the portfolio to be maximally diversified.

III Applications

In this section, the optimization algorithm described above is applied in two settings, in order to demonstrate the potential value of the new procedure. First, least absolute deviations (LAD) and least squares (LS) methods are applied to international stock returns data; it is seen that the LAD estimator leads to superior out–of–sample predictions. Next, the branch–and–bound integer programming method is used to create an optimal stock index based on assets trading New York Stock Exchange; here, it is shown that an accurate index can be formed from a relatively small number of assets.

International Stock Returns

Data for this study were obtained from the DRI Basic Economics database, and consisted of monthly nominal returns on stock indices from Canada, France, Italy, Japan, United Kingdom, United States, and West Germany, from 1950 through 1995. The West Germany (W.G.) return series was chosen as the target portfolio m_t, and the remaining countries' stock returns were used for the y_{jt}. The goal of the analysis was to select portfolios of non–W.G. returns which could accurately match, or hedge, the W.G. market return. Two methods were used: an LAD estimator, and an LS estimator. In both cases, the number of assets receiving non-zero allocations was unconstrained.

22 datasets of monthly returns were obtained, each of length 48 months. The first dataset contained data from 1950–53, the next from 1952-1955, and the last from 1992-1995. For each dataset, the first 24 months of data were used to form the optimal portfolios. The remaining 24 months of data were used to evaluate the ability of the constructed portfolio p_t to match the W.G. portfolio m_t out-of-sample. The out-of-sample accuracy measures used for the study were the mean absolute error (MAE), $\frac{1}{24}\sum_{t=25}^{48}|m_t - p_t|$, and the mean square error (MSE), $\frac{1}{24}\sum_{t=25}^{48}(m_t - p_t)^2$; these measures were computed for each of the 22 consecutive datasets.

In terms of MAE, the LAD method outperformed the LS method in 15 out of 22 periods (p–value = 0.026, binomial sign test), and in terms of MSE, the LAD method outperformed the LS method in 17 out of 22 periods (p–value = 0.002, binomial sign test). The median gain in accuracy of the LAD method relative to the LS method, over the 22 test periods, was 9% for the MAE measure (standard deviation=12%), and 15% for the MSE measure (standard deviation=22%). The LAD method appears to have an advantage in analyzing the long–tailed returns data.

N.Y.S.E. Data

Data for this study were obtained from the Center for Research in Stock Prices (C.R.S.P.) dataset. The target portfolio m_t was chosen as the C.R.S.P. Value Weighted Index (V.W.I.), which is a measure of overall stock market performance. The component assets used for the y_{jt} were fifty randomly selected stocks from the C.R.S.P. database; all these stocks trade on the New York Stock Exchange (N.Y.S.E.). The goal of the analysis was to select portfolios, based on subsets of the fifty stocks, which could accurately match the V.W.I. market return, using the integer programming method of Section II to constrain the number of securities entering the index portfolio.

19 datasets of monthly returns were obtained, each of length 48 months. The first dataset contained data from 1955–58, the next from 1957–1960, and the last from 1991–1994. For each dataset, the first 24 months of data were used to form the optimal portfolio, using a branch–and–bound algorithm to solve optimization problem (2), with minimand $\frac{1}{24}\sum_{t=1}^{24}|m_t - p_t|$. The remaining 24 months of data were used to evaluate the ability of the constructed portfolio p_t to match the V.W.I. portfolio m_t out–of–sample; the out-of-sample accuracy measure used was the mean absolute error, $\frac{1}{24}\sum_{t=25}^{48}|m_t - p_t|$.

A number of different methods were used to form the indexing portfolio p_t:

50e – an equally weighted portfolio using all 50 stocks

20o – a portfolio with at most 20 stocks, with the 20 stocks and the weights chosen optimally

20r – a portfolio with 20 stocks selected at random, but with weights chosen optimally

5o – a portfolio with at most 5 stocks, with the 5 stocks and the weights chosen optimally

5r – a portfolio with 5 stocks selected at random, but with weights chosen optimally

Table 1 displays the results of the analysis. It can be seen from the table that optimization can indeed lead to accurate tracking of a target out–of–sample. Portfolio **5o**, which optimally selects the five assets to receive positive weight, as well as the appropriate weights for the assets, performs better at tracking the overall market measure V.W.I. than does portfolio

5r, in which the five assets are selected at random; the difference is statistically significant ($p = .031$, binomial sign test). The difference in performance between portfolios **20o** and **20r** is statistically significant as well ($p = .0096$). The optimized portfolio with five assets, **5o** outperformed the equally weighted portfolio with fifty assets, **50e**, indicating the potential economy that can be achieved with optimized indices. The average computing time required to identify the **5o** portfolios was less than 0.2 CPU seconds, using a Sun Sparc 10 Workstation with 150 MHz processor.

[Insert Table 1 Here]

IV Discussion

The mathematical theory of portfolio selection in the face of significant, fixed transaction costs has been treated, for example, in Akian, Menaldi, and Sulem (1995) and Cvitanic and Karatzas (1996). The mixed integer programming framework presented in Section II of this paper provides a practical algorithm for selecting a hedging or indexing portfolio whose historical returns optimally match those of a target portfolio, while keeping the number of assets included in the matching portfolio small. The empirical example of the N.Y.S.E. portfolio presented in Section III shows that index optimization can successfully produce accurate tracking of a target with a small number of intelligently chosen assets.

Kariya (1993, chapter 12) discusses alternative approaches for choosing an index portfolio. These approaches also aim to achieve close matching of a target portfolio, with a small number of assets. The approach advocated in Kariya (1993) differs from the method introduced in this paper in a few ways. First, Kariya (1993) uses an L_2, or least sum of squares, objective criterion, whereas the present article uses an L_1, or least absolute deviations criterion. The LAD criterion will have advantages in the presence of outliers in the data; For returns data that are approximately normally distributed, though, the LS criterion could lead to more efficient estimates of the optimal portfolio weights. Another difference between the method of Kariya (1993) and the present paper is in the method by which the subset of assets to be included is determined. The present paper uses global optimization, via branch–and–bound methods, to choose the subset which leads to optimal tracking in a training sample. Kariya (1993) uses a heuristic method which involves the use of cluster analysis to identify assets that behave similarly, in the training sample, to the target being tracked.

The least absolute deviations method for measuring riskiness of financial portfolios was used in Yamakazi and Konno (1991). Other related papers are Stone (1973) and Sharpe (1971), which discuss approximate methods for minimizing portfolio volatility – measured in the L_2 norm – via linear programming methods. Young (1998) discusses the use of a minimax, rather than least absolute deviations, criterion in portfolio optimization, and shows how this criterion can also be treated within a linear or linear–integer programming framework.

Given a specified subset of assets to be included in the index or hedge portfolio, the optimal weights for the assets are obtained via a simple linear programming problem. Perhaps the more challenging task in finding an optimal hedging portfolio is the selection of the subset of assets to receive positive portfolio allocations. The branch–and–bound method used in this paper works quite efficiently for moderate to large sized problems. Future research in the area of indexed or hedge portfolio optimization may involve the use of alternative optimization techniques - e.g., genetic algorithms (Michalewicz and Janikow 1991) or tabu search (Glover 1994) – to explore the space of candidate asset subsets.

Figure 1: Scatter plot of monthly stock returns, 1991–1992, with estimated regression lines: Solid = LAD, Dotted = LS

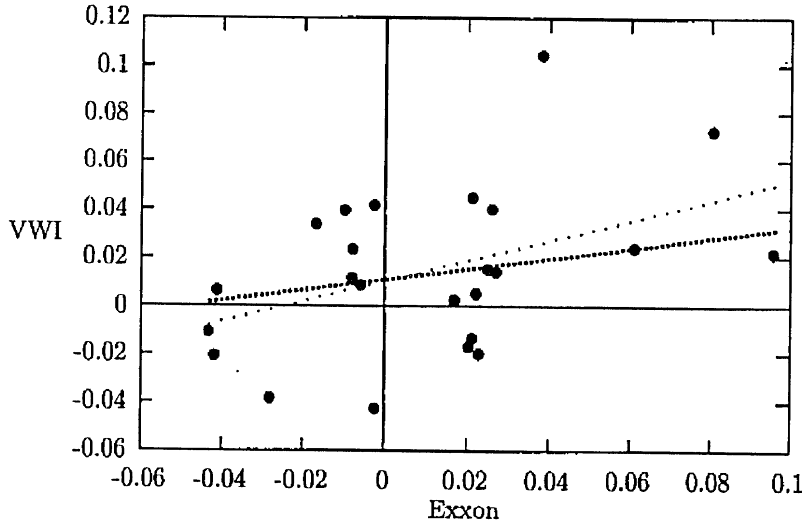

Table 1: Results of N.Y.S.E. data analysis. MAE denotes out–of–sample mean absolute error; Median and SD denote the summary statistics of the MAE over the 19 periods of analysis.

	MAE	
Method	Median	SD
50e	0.0787	0.0249
5	0.0735	0.0301
5c	0.0927	0.0297
20	0.0734	0.0312
20c	0.0833	0.0321

References

Akian, M., J.-L. Menaldi, and A. Sulem (1995). Multi-asset portfolio selection problem with transaction costs. *Mathematics and Computers in Simulation 38*, 163–172.

Anderson, R. W. and J. P. Danthine (1981). Cross hedging. *Journal of Political Economy 89*, 1182–1196.

Armstrong, R. D. and J. Godfrey (1979). Two linear programming algorithms for the linear discrete L_1 norm problem. *Mathematics of Computation 33*, 289–300.

Barrodale, I. and F. Roberts (1978). An efficient algorithm for discrete L_1 linear approximation. *SIAM Journal of Numerical Analysis 15*, 603–611.

Cvitanic, J. and I. Karatzas (1996). Hedging and portfolio optimization under transaction costs: a martingale approach. *Mathematical Finance 6*, 133–165.

Dielman, T. and R. Pfaffenberger (1982). LAV (least absolute value) estimation in linear regression: A review. *TIMS/Studies in the Management Sciences 19*, 31–52.

Fama, E. F. (1965). The behavior of stock prices. *Journal of Business 38*, 34–105.

Glover, F. (1994). Tabu search for nonlinear and parametric optimization (with links to genetic algorithms). *Discrete Applied Mathematics 49*, 231–255.

Gonin, R. and J. Money (1989). *Non-linear L_p-Norm Estimation.* New York, NY: John Wiley and Sons.

Herbst, A. F. and J. F. Marshall (1995). Convergence-adjusted composite hedging. *The Journal of Financial Engineering 3*, 109–135.

Hull, J. (1997). *Options, Futures, and Other Derivatives* (3rd ed.). Upper Saddle River: Prentice–Hall, Inc.

Kariya, T. (1993). *Quantitative Methods for Portfolio Analysis: MTV Model Approach.* Norwell, MA: Kluwer Academic Publishers.

Koenker, R. and S. Portnoy (1997). Quantile regression. Working Paper 97-0100, University of Illinois College of Commerce and Business Administration.

Kon, S. J. (1984). Models of stock returns – a comparison. *Journal of Finance 39*, 147–165.

Lehmann, E. L. (1983). *Theory of Point Estimation.* New York, NY: John Wiley and Sons.

Lucena, A. and J. E. Beasley (1996). Branch and cut algorithms. In *Advances in Linear and Integer Programming*, pp. 187–221. New York, NY: Oxford University Press.

Markowitz, H. M. (1952). Portfolio Selection. *Journal of Finance 7*, 77–91.

McCulloch, J. H. (1978). Continuous time processes with stable increments. *Journal of Business 51*, 601–619.

McKeown, P. G. (1981). A branch-and-bound algorithm for solving fixed charge problems. *Naval Research Logistics Quarterly 28*, 607–617.

Michalewicz, Z. and C. Z. Janikow (1991). Genetic algorithms for numerical optimization. *Statistics and Computing 1*, 75–91.

Sharpe, W. (1971). A linear programming approximation for the general portfolio analysis problem. *Journal of Financial and Quantitative Analysis 6*, 1263–1275.

Stone, B. (1973). A linear programming formulation of the general portfolio selection model. *Journal of Financial and Quantitative Analysis 8*, 621–636.

Suhl, U. H. (1985). Solving large scale mixed-integer programs with fixed charge variables. *Mathematical Programming 32*, 165–182.

Yamakazi, H. and H. Konno (1991). Mean absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science 37*, 519–531.

Young, M. R. (1998). A minimax portfolio selection rule with linear programming solution. *Management Science*.