On the Optimal Selection of Portfolios under Limited Diversification

Jay Sankaran
Department of Management Science and Information Systems
University of Auckland
New Zealand
j.sankaran@auckland.ac.nz

C. Krishnamurti
Department of Finance and Accounting
National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260.

Ajay A. Patil
Computervision Corporation
Bedford, MA 01730, U.S.A.

Abstract

We address the problem of selecting portfolios that are optimal among all those portfolios that comprise at most a pre-specified number, *k*, of securities. We consider two criteria: maximizing the ratio of the average excess return to the standard deviation; and maximizing the correlation with a specified market-index. Under standard assumptions of the behaviour of stock returns, we develop procedures that are of polynomial complexity and that require minimal data.

---

1 Introduction

One basic implication of modern portfolio theory is that investors hold well-diversified portfolios. However, there is empirical evidence that individual investors typically hold only a small number of securities. Blume, Crockett, and Friend [7] found that only 11% of individual U.S. investors held more than ten stocks. Conine, Jensen, and Tamarkin [9, footnote 2, pp. 1004-1005] cite several studies that provide empirical evidence that the majority of individual investors in the U.S. hold highly undiversified portfolios. Bark [5] states that one reason for the inadequacy of the Sharpe-Lintner-Mossin capital asset pricing model in the Korean stock market is that the portfolios of Korean investors are also highly undiversified.

The absence of diversification has also been observed with regard to index-tracking portfolios. Rudd [26] provides evidence which indicates that the majority of American index funds do not actually hold all the stocks in the chosen index - some of them hold as little as 35 stocks to match the 500-stock S&P 500 index.

Market imperfections such as fixed transaction costs provide one explanation for the prevalence of undiversified portfolios. Indirect transaction costs, such as the cost of analyzing securities, also militate against diversification; a small investor who chooses to invest in only a limited number of securities can devote more attention to the individual behavior of those securities and their mean-variance characteristics. Moreover, there is evidence [16, 18, 19] that diversification beyond 8-10 securities may not be worthwhile provided these securities are chosen not randomly but through a systematic, optimum-seeking procedure. Citing Szegö [30], Sengupta and Sfeir [27] observe that it may be superfluous to enlarge the number of securities in a portfolio beyond a limit because the variance-covariance matrix of the returns on the securities in a portfolio that has a large number of securities tends to conceal significant singularities or near-singularities.

2 Problem Definition

We consider two related problems in optimal portfolio selection. The first, referred to as P1, is to find portfolios that are mean-variance efficient and that comprise at most a pre-specified number of securities. The second problem, referred to as P2, is to find portfolios that optimally track a specified market-index among all those portfolios that contain at most a pre-specified number of securities and whose average returns equal the average return on the index. We determine the efficiency of tracking by measuring the correlation coefficient between the returns on the chosen market index and the returns of the market-tracking portfolio.

3 Existing Algorithms for P1 and P2 and Variants Thereof

Mao [22] and Jacob [19] formally address P1 but develop their selection procedures under somewhat restrictive assumptions and rather high degrees of approximation. For instance, to compute the number of securities that optimally trades-off diversification against (fixed) transaction costs, Mao assumes that both the average excess return over the riskless rate and the standard deviation of the return are the same for all the securities in the portfolio. Further, for selecting the best portfolio among those that comprise a pre-specified number of securities, he assumes that for all of these portfolios, the nonsystematic risk is fully diversified away. Jacob assumes that the weights for the securities in the investor's portfolio are all equal to each other. She also linearizes portfolio risk in terms of the weights.

For the problem of determining mean-variance efficient portfolios under the single-index model [28] of security returns and an upper limit on the number of stocks, Faaland [17] develops an algorithm based on integer programming, which is bettered by the implicit enumeration algorithm of Blog et al. [6]. Cooper and Farhangian [11] develop a dynamic programming approach for an extension of this problem that incorporates fixed costs of transaction.

Assuming that the capital asset pricing model [21, 24, 29] holds, Brennan [8] presents an algorithm for determining the optimal number of securities under fixed transaction costs. However, the validity of that assumption in the presence of fixed transaction costs has been questioned [25]. Patel and Subrahmanyam [25] develop an efficient algorithm for the problem under the assumption that the correlation coefficient is the same for all pairs of securities [12, pp. 168-169]. Aneja, Chandra, and Gunay [4] show how the average pairwise correlation coefficient can be efficiently estimated using a portfolio approach.

We are not aware of more recent algorithms for P1 or any of its variants. Nevertheless, simple ranking procedures akin to the well-known EGP algorithms [13, 14, 15] continue to be developed for selecting mean-variance efficient portfolios in *other* contexts such as restricted short-selling [2, 3] and institutional norms for short-selling [20].

Procedures for P2 or variants thereof have been propounded by Adcock and Meade [1], Connor and Leland [10], Meade and Salkin [23], and Rudd [26] among others; however, they use optimization only to the extent of determining the portfolio weights for a pre-selected set of securities. Thus, for instance, in one of their procedures, Meade and Salkin [23] randomly sample a small number of stocks, and then use quadratic programming to determine the optimal portfolio weights.

4 The Results

To formally present the problems and algorithms, we employ the following notation.

n: The number of securities in the universe.

N: The set of securities in the universe; thus, $N = \{1,\ldots,n\}$.

k: A pre-specified upper limit on the number of securities in the portfolio ($1 \le k \le n$).

x_i: The weight of security i, i = 1,...,n.

R: The rate of return on the riskless asset.

R_i: The expected rate of return on security i, i = 1,...,n.

s_i: The standard deviation of the rate of return on security i, i = 1,...,n.

b_i: The ratio of the average excess return to the standard deviation of security i, i = 1,...,n; thus, $b_i = (R_i - R)/s_i$.

p: An estimate of the average correlation coefficient of any pair of security returns (we assume that p is non-negative). The assumption of non-negativity for p would appear to be very mild because a negative value of p would imply that the variance of an equally-weighted portfolio with, say, m securities, is negative for sufficiently large values of m (see, for instance, [12, p. 60]).

4.1 Problem P1

Under the assumption of constant pairwise correlations, P1 may be formulated as:

$$\text{Maximize} \frac{\sum_{i=1}^{n} (R_i - R).x_i}{\sqrt{\sum_{i=1}^{n} s_i^2 x_i^2 + p.\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} s_i s_j x_i x_j}}$$

such that

at most *k* of $\{x_i$: i = 1,...,n$\}$ are non-zero, and

if short-selling is disallowed, then $x_i \geq 0$ for i = 1,...,n.

In the maximand, the numerator is the excess return on the portfolio while the denominator is the standard deviation of the return on the portfolio; by varying the riskless rate of return, R, we can generate the mean-variance-efficient frontier. We assume that for any given value of R, the *n* stocks in the universe are numbered in descending order of $\{b_i$: i = 1,…,n$\}$.

If short-selling is allowed, then the set of stocks to select is of the form {1,…,t, n-k+t+1, ..., n}, where $0 \leq t \leq k$. Thus, the correct algorithm is to evaluate the k+1 portfolios and choose the best among them.

If short-selling is not allowed, then the correct algorithm is, quite simply, to execute the well-known EGP ranking algorithm until *k* securities are included. We also establish that the optimal objective value of the above problem increases with *k*, but at a decreasing rate; in other words, the marginal benefit from diversification decreases with the number of stocks in the portfolio.

4.2 Problem P2

The problem, P2, is to find a portfolio:

(i) which uses at most *k* securities from the universe of n stocks;

(ii) whose expected return equals that on the given market-index; and

(iii) whose return has the greatest correlation with that on the index, among all portfolios that satisfy conditions (i) and (ii).

For formalizing the problem and presenting the results, we introduce some additional notation:

X_i: The investment (not the weight) in security i.

π_i: The correlation of the return on security i with that on the market-index.

s_p: The standard deviation of portfolio return.

Assuming the presence of a riskless asset, P2 can be transformed to:

$$\text{Min } (s_p)^2 \text{ such that}$$
$$\text{at most } k \text{ from among } \{X_i: i = 1,...,n\} \text{ are non-zero, and}$$
$$\text{if short-selling is disallowed, all of } \{X_i: i = 1,...,n\} \text{ are non-negative.}$$

We assume that stocks are numbered in descending order of correlation with the index.

Under the single-index model of stock returns, the correct algorithm is to pick stocks in descending order of correlation (if short-selling is allowed, we use the absolute values of the correlations) until the limit, *k*, is reached. We can readily establish that regardless of whether or not short-selling is allowed, the benefit from diversification decreases with the pre-specified upper limit on the number of stocks in the portfolio.

Under the assumption of constant pairwise correlations, the following results apply.

(a) If short-selling is allowed, then the set of stocks to select is of the form {1,…,t, n-k+t+1, ..., n}, where $0 \leq t \leq k$. Thus, the correct algorithm is to evaluate the k+1 portfolios and choose the best among them.

(b) If short-selling is disallowed, the correct procedure is to execute the EGP algorithm (on $\{\pi_i: i = 1,\dots,n\}$ rather than $\{b_i: i = 1,\dots,n\}$) until k securities are included. In this case, we can establish that diversification yields diminishing returns (no pun intended).

A potential problem in applying the above procedures is that under constant pairwise correlations, the estimated (*n*+1)x(*n*+1) variance-covariance matrix (involving the *n* securities and the market-index) may not be positive definite.

5 Summary of Findings

We have presented algorithms for selecting `small' portfolios (namely, those that contain at most a pre-specified number of securities) under two different objective criteria. The first is to maximize the ratio of excess return to standard deviation (and thereby, generate the mean-variance-efficient frontier), while the second is to maximize the correlation with a given market-index – the corresponding problems are referred to as P1 and P2 respectively.

Under the assumption of constant pairwise correlations, exact and efficient (polynomial) procedures exist for both P1 and P2 regardless of whether or not short-selling is allowed. Further, P2 can also be efficiently solved under the single-index model of stock returns for both the cases (i.e., short-selling allowed and short-selling disallowed).

It is important to emphasize that the requirements in terms of both data and computing power are minimal for the algorithms presented herein. Moreover, empirical results on data from the Korean stock exchange suggest that the optimal portfolio weights are neither too large nor too small.

Future research could explore problems P1 and P2 under more general correlation structures. Under the single-index assumption, P1 is NP-hard; thus, the single-index case might need enumeration. At least in the case when short-selling is disallowed, both lower and upper bounds on the optimal objective value of any node in the enumeration tree can be obtained through the results of the paper.

References

[1] C.J. Adcock, N. Meade, *A Simple Algorithm to Incorporate Transactions Costs in Quadratic Optimisation*, European Journal of Operational Research, 79 (1994), pp 85--94.

[2] G.J. Alexander, *Short Selling and Efficient Sets*, Journal of Finance, 48 (1993), pp 1497--1506.

[3] G.J. Alexander, *Efficient Sets, Short-Selling, and Estimation Risk*, Journal of Portfolio Management, 21 (1995), pp 64--73.

[4] Y.P. Aneja, R. Chandra, E.A. Gunay, *A Portfolio Approach to Estimating the Average Correlation Coefficient for the Constant Correlation Model*, Journal of Finance, 44 (1989), pp 1435--1438.

[5] H.K. Bark, *Risk, Return, and Equilibrium in the Emerging Markets: Evidence from the Korean Stock Market*, Journal of Economics and Business, 43 (1991), pp 353--362.

[6] B. Blog, G. van der Hoek, A.H.G. Rinnooy Kan, G.T. Timmer, *The Optimal Selection of Small Portfolios*, Management Science, 29 (1983), pp 792--798.

[7] M.E. Blume, J. Crockett, I. Friend, *Stock Ownership in the United States: Characteristics and Trends*, Survey of Current Business, 54 (1974), pp 16--40.

[8] M.J. Brennan, *The Optimal Number of Securities in a Risky Asset Portfolio when there are Fixed Costs of Transacting: Theory and some Empirical Results*, Journal of Financial and Quantitative Analysis, 10 (1975), pp 483--496.

[9] T.E. Conine, Jr., O.W. Jensen, M. Tamarkin, *On Optimal Production and the Market to Book Ratio given Limited Shareholder Diversification*, Management Science, 35 (1989), pp 1004--1013.

[10] G. Connor, H. Leland, *Cash Management for Index Tracking*, Financial Analysts Journal, 51 (1995), pp 75--80.

[11] M.W. Cooper, K. Farhangian, *An Integer Programming Algorithm for Portfolio Selection with Fixed Charges*, Naval Research Logistics Quarterly, 29 (1982), pp 147--150.

[12] E.J. Elton, M.J. Gruber, *Modern Portfolio Theory and Investment Analysis* (fifth edition), John Wiley & Sons, New York (1995).

[13] E.J. Elton, M.J. Gruber, M.W. Padberg, *Simple Criteria for Optimal Portfolio Selection*, Journal of Finance, 31 (1976), pp 1341--1357.

[14] E.J. Elton, M.J. Gruber, M.W. Padberg, *Simple Rules for Optimal Portfolio Selection: The Multi Group Case*, Journal of Finance, 32 (1977), pp 329--345.

[15] E.J. Elton, M.J. Gruber, M.W. Padberg, *Simple Criteria for Optimal Portfolio Selection: Tracing out the Efficient Frontier*, Journal of Finance, 33 (1978), pp 296--302.

[16] J.L. Evans, S.H. Archer, *Diversification and the Reduction of Dispersion: An Empirical Analysis*, Journal of Finance, 29 (1968), pp 761--767.

[17] B.H. Faaland, *An Integer Programming Algorithm for Portfolio Selection*, Management Science, 20 (1974), pp 1376--1384.

[18] L. Fisher, J.H. Lorie, *Some Studies of the Variability of Returns on Investments in Common Stocks*, Journal of Business, XLIII (1970), pp 99--134.

[19] N. Jacob, *A Limited-Diversification Portfolio Selection Model for the Small Investor*, Journal of Finance, 29 (1974), pp 847--856.

[20] C.C.Y. Kwan, *Optimal Portfolio Selection under Institutional Procedures for Short Selling*, Journal of Banking & Finance, 19 (1995), pp 871--889.

[21] J. Lintner, *The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets*, Review of Economics and Statistics, 51 (1965), pp 13--37.

[22] J.C.T. Mao, *Essentials of Portfolio Diversification Strategy*, Journal of Finance, 25 (1970), pp 1109--1131.

[23] N. Meade, G.R. Salkin, *Index Funds - Construction and Performance Measurement*, Journal of the Operational Research Society, 40 (1989), pp 871--879.

[24] J. Mossin, *Equilibrium in a Capital Asset Market*, Econometrica, 34 (1966), pp 768--783.

[25] N.R. Patel, M.G. Subrahmanyam, *A Simple Algorithm for Optimal Portfolio Selection with Fixed Transaction Costs*, Management Science, 28 (1982), pp 303--314.

[26] A. Rudd, *Optimal Selection of Passive Portfolios*, Financial Management, 9 (1980), pp 57--66.

[27] J.K. Sengupta, R.E. Sfeir, *Tests of Efficiency of Limited Diversification Portfolios*, Applied Economics, 17 (1985), pp 933--945.

[28] W.F. Sharpe, *A Simplified Model for Portfolio Analysis*, Management Science, 9 (1963), pp 277--293.

[29] W.F. Sharpe, *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*, Journal of Finance, 19 (1964), pp 425--442.

[30] G.P. Szegö, *Portfolio Theory*, Academic Press, New York (1980).