Proceedings of the World Congress on Engineering and Computer Science 2011 Vol II
WCECS 2011, October 19-21, 2011, San Francisco, USA

Using MOEAs to Outperform Stock Benchmarks in the Presence of Typical Investment Constraints

Andrew Clark and Jeff Kenyon

Abstract—**Portfolio managers are often constrained by turnover limits, minimum and maximum stock positions, cardinality, a target market capitalization and sometimes the need to hew to a style (growth or value). In addition, many portfolio managers choose stocks based upon fundamental data, e.g. price-to-earnings and dividend yield in an effort to maximize return. All of these are typical real-world constraints a portfolio manager faces. Another constraint, of sorts, is the need to outperform a stock index benchmark. Performance higher than the benchmark means a better compensation package. Underperforming the benchmark means a lesser compensation package.**

We use MOEAs to satisfy the above real-world constraints and consistently outperform typical performance benchmarks. Our first MOEA solves all the constraints (except turnover and position limits) and generates feasible portfolios. The second MOEA tests each of the potential feasible portfolios of the first MOEA by trading off mean return, variance, turnover and position limits. The best portfolio is chosen from these feasible portfolios and becomes the portfolio of choice for the next quarter.

The MOEAs are applied to the following problems - generate a series of monthly portfolios that outperform the S&P 500 over the past 30 years and generate a set of monthly portfolios that outperform the Russell 3000 Growth index over the last 15 years. Our two MOEAs accomplish both these goals on a risk adjusted and non-risk adjusted return basis.

Index Terms—**multi-objective evolutionary algorithms (MOEA), mean-variance optimization, financial constraints, multi-period MOEAs**

I. Introduction

IN finance, a portfolio is a collection of assets held by an institution or a private individual. The portfolio selection problem seeks the optimal way to distribute a given monetary budget on a set of available assets. The problem usually has two criteria: maximizing return and minimizing risk. Classical mean-variance portfolio selection aims at simultaneously maximizing the expected return of the portfolio and minimizing portfolio risk. In the case of linear constraints, the problem can be solved efficiently by parametric quadratic programming (i.e., variants of Markowitz' critical line algorithm). What complicates this simple statement of portfolio construction are real-world constraints that are by definition non-convex, e.g. cardinality constraints which limit the number of different assets in a portfolio and minimum buy-in thresholds. In what follows, we solve the constrained portfolio construction problem with two multi-objective evolutionary algorithms (MOEAs). The idea is to let the first MOEA come up with the set of all feasible portfolios and have the second MOEA generate the efficient frontier for each feasible portfolio. The best solution from the second MOEA is determined using the standard distance function to measure dominance in the MOEA solution sets.

Manuscript received June 22, 2011; revised July 17, 2011.

Andrew Clark is Chief Index Strategist for Thomson Reuters Indices (e-mail: andrew.clark@thomsonreuters.com).

Jeff Kenyon is Lead Software Engineer for Thomson Reuters Indices (e-mail: jeff.kenyon@thomsonreuters.com).

II. Problem Statement and Pseudo-code

The two problems we will work with are: generate a series of monthly portfolios that outperform the S&P 500 over the last 30 years and generate a set of monthly portfolios that outperform the Russell 3000 Growth index over the last 15 years.

The common constraints we will operate under are: turnover not to exceed 8% per month, a minimum stock position set at 0.35% of the net asset value of the portfolio, a maximum stock position of 4% of the net asset value of the portfolio[1], we must choose stocks that maximize the scores generated by a multi-factor stock model[2] (this constraint typifies the use of what is called fundamental financial data to select stocks that are potential candidates for the final monthly portfolios) and finally a target market capitalization constraint where the average market capitalization of the portfolio must be greater than the average market capitalization of all stocks available to purchase in the current month (this last constraint will mean the portfolios must be what is called "large-cap." Both S&P 500 and the Russell 3000 Growth are large-cap indices [portfolios]).

An additional constraint that we will use just for the Russell 3000 Growth problem is: we cannot exceed the average book-to-price value of all stocks available for purchase in the current month. Meeting this constraint means we will generate the required growth portfolios in order to conform to the style of the Russell 3000 Growth.

We solve the issues of the constraints by breaking them into two groups and using two MOEAs. The first MOEA generates potential portfolios that lie within the bounds of all the constraints except turnover and position constraints. In the second MOEA, we trade off the turnover and position constraints against mean return and variance (the last being the typical factors used in mean-variance optimization). We set the rebalance period to be quarterly versus monthly but stay within the stated turnover constraint (not to exceed 8% per month). The steps each MOEA follows are next.

[1] The minimum and maximum position constraints also imply a cardinality constraint. Dividing 100% by 0.35% yields 286 which is the maximum number of stocks any portfolio can have. Dividing 100% by 4% gives 25 which the minimum number of stocks any portfolio can have. The cardinality constraint is solved in the first MOEA while the position constraints are solved in the second MOEA.

[2] The details and back testing of the multi-factor stock model can be found in [1]. We thank John for supplying the multi-factor scores on a stock-by-stock basis for the past 30 years.

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

A. Data Loading, Retrieving and Filtering

1) Load the target candidate constituent sets from the multi-factor score files. "Sets" and "Files" because we need to build two sets of candidate portfolios each month, one for the large-cap portfolio and another for the large-cap growth portfolio (please note that the size of these files goes from approximately 1000 stocks in 1980 to more than 3000 in 2009).
2) Remove candidates that may be excluded, on an *a priori* basis, of being unable to contribute toward the portfolio goals.
    a) Remove candidates with scores of less than 20 (this was a heuristic choice)
    b) Rank candidates by market cap and eliminate the bottom 12% of candidates. If the market cap at this level is greater than US$750M, use US$750M as the floor for the cutoff (this cutoff is based upon common definitions of where small-cap stocks start to appear the U.S. stock market).
3) From this subset of candidate equities, retrieve daily price data going back 287 observations (for use in trading off mean and variance in second MOEA), and going forward 63 observations (for use in performance calculation of final [best] portfolio over the next three months).
4) Calculate the average market cap and average book to price score of the constituents in the target benchmark index (average book-to-price is used to determine if a portfolio satisfies the style constraint).
5) The candidate constituents remaining serve as the primary input to the MOEA phase.

B. MOEA Phase I

In this phase the goal is to identify a set of portfolios (ideally 50 or less) that will be examined by the second MOEA. Output of the MOEA for each portfolio is the identification of a subset of the candidates (between 25 and 286) to be used as candidates for later optimization (again 25 because of the 4% maximum constraint and 286 because of 0.35% constraint).

1) The MOEA is invoked, passing the candidate constituent data, portfolio constituents from previous rebalances, the market cap average for the target benchmark index and population, generations, and mutation rate. For the large-cap growth portfolio, there is an additional parameter for the average book-price score of the target benchmark portfolio.
    a) MOEA algorithm is NSGA II, using single point crossover, bit flip mutation, and binary tournament for selection. For the passed parameters, population represents the number of proposed solutions that will be carried from generation to generation (note that the number of non-dominated solutions is often considerably smaller); generations is the number of generations to be evaluated within the evolutionary algorithm; and mutation rate indicates the rate at which a dominated portfolio will be modified as it moves from generation to generation.
    b) For the large-cap portfolio, population is 500, generations are 1200, and mutation rate is 0.03 (3%).
    c) For large-cap growth portfolio, population is 50 (for the three-objective problem, a population of 500 tended to generate several hundred non-dominated solutions; the population number is dropped to avoid generating more solutions than can be explored in the MVO phase). All other parameters are the same.
2) Generation-zero portfolios are started with 156 equities randomly selected. If there are portfolios from a previous rebalance available, those portfolios are seeded into the population.
3) The large-cap portfolio objectives are the maximization of multi-factor score and average market cap. The large-cap growth portfolio adds a third objective to minimize the book-to-price score.
4) Penalties are in place to enforce the cardinality and market cap constraints. For large-cap growth, there is also a penalty for exceeding the book-price average passed in.
5) The MOEA phase produces two files as output. One file contains the objective values for the solution set, and the other contains the portfolios (where each equity has a 0/1 value) describing the Pareto efficient frontier.
6) Any portfolios passing constraints from the previous rebalance are added to the set of portfolios generated by the MOEA.

C. MOEA Phase II

Some definitions are needed to understand the workings of the second MOEA.

First is the efficient frontier. This frontier is calculated by trading off mean stock returns and their related variances. In essence, a combination of stocks (the portfolio) is called efficient if it has the best possible expected level of return for its level of risk (this risk is usually usually proxied by the standard deviation of the portfolio's return). Every possible combination of stocks can be plotted in risk-expected return space and the collection of all such possible portfolios defines a region in this space. The upward-sloping boundary of this region, a hyperbola, is called the efficient frontier.

The Sharpe ratio is a measure of the excess return (or risk premium) per unit of risk in a portfolio or a trading strategy. The Sharpe ratio is the standard measure of the risk premium when an efficient frontier is calculated. The Sharpe ratio is used to characterize how well the return of a portfolio compensates the investor for the risk taken. The higher the Sharpe ratio the better the portfolio trades off risk and return.

When comparing portfolios with differing expected returns against the risk-free rate (in our case 3-month U.S. Treasury bills), the portfolio with the higher Sharpe ratio gives more return for the same risk. Investors are often advised to pick portfolios with high Sharpe ratios. The best Sharpe ratio on the efficient frontier is by definition the best portfolio to invest in.

This short introduction to a part of modern portfolio theory (MPT) can be supplemented by [2].

Returning to the second MOEA, daily returns are calculated for all equities that may appear in a portfolio (to avoid having to recalculate a daily return series for an equity multiple times), and the daily risk free rate closest to the rebalance date is retrieved. Then for each feasible portfolio from the previous MOEA, the following actions are taken:

1) Identify the equities designated as candidates for the portfolio being processed, and form the returns matrix. The minimum number of returns to be used is 126 (six months); because the number of observations must exceed the number of candidate equities (the so-called "curse of dimensionality"), the maximum number of returns in the series is determined by portfolio size, up to 287.
2) In addition to forming the returns matrix (step 1, above), we also need to generate mean expected returns, variances, and the covariance matrix.
3) The MOEA is called, using pointers to the above files, a pointer to the file containing the previous winning portfolio, and the MOEA parameters (population = 100, generations = 600, mutation rate = 0.01).
4) Following the lead of an earlier replication of MVO results using MOEA, the algorithm used is SPEA2. SBXCrossover, polynomial mutation, and selection by binary tournament are used. The SPEA2 archive size is the same as the population size.
5) Random portfolios are used for generation zero. If available, the previous winning portfolio is seeded into the generation zero set.
6) In each evaluation, the randomly assigned weights are normalized (so that they sum to 1).
7) In weighting strategy #1 (non-zero values are $0.0035 \leq w \leq 0.04\%$), weights greater than 0.00175 are rounded up to 0.0035, while weights under that amount are rounded to zero. Weights over 0.04 are set to 0.04. All weight adjustments are added to a ledger, and then debited or credited at the end of the adjustment process (evenly divided among those that can accept the debit/credit amount without moving outside the set weight boundary).
8) In weighting strategy #2 (all equities are weighted, $0.0035 \leq w \leq 0.04\%$), values less than 0.0035 are rounded up to 0.0035, while values over 0.04 are set to 0.04. All weight adjustments are added to a ledger, and then debited or credited at the end of the adjustment process (evenly divided among those that can accept the debit/credit amount without moving outside the set weight boundary).
9) There are four objectives for the MOEA: maximize return, minimize risk, minimize turnover and meet the maximum and minimum holdings constraint. The first two objectives are the standard MVO calculations [3]. The turnover and position objectives calculate the shift in weight between the previous winning portfolio and the proposed portfolio. It attempts to insure that no portfolios break the holdings or turnover bounds.
10) Once all MVOs have been run, the winning portfolio is then selected. Ideally, this is the portfolio with the best Sharpe ratio that has also passed the market cap and turnover constraints. If no portfolios pass the constraints, the portfolio with the best turnover is considered the winner.
11) The performance for the winning portfolio (and the target benchmark index) for the period until the next rebalance is then calculated and logged.

D. *MOEA Phase IIa*

If the turnover constraint is not met in Step 9, we use a subsidiary MOEA that lies inside the second MOEA which trades off the best Sharpe ratio portfolio with its existing stock positions against turnover.

The form this MOEA takes is somewhat similar to solving the minimum cut problem in graph theory using MOEAs (see, for example, [4]). The difference between our MOEA and the MOEAs that have been used to solve the minimum cut problem is that we are interested in reducing or enlarging the weight of one or more edges while keeping within the position limit constraints (though we do allow stock holdings to go to zero if needed). Once stocks are re-weighted and the turnover constraint reached, a new efficient frontier is calculated and the new Sharpe portfolio examined. If the new Sharpe portfolio meets the turnover and position constraints, the MOEA in IIa stops and the best Sharpe portfolio for this feasible portfolio from MOEA I is stored. The process is repeated for all the portfolios passed from MOEA I that do not meet the turnover constraints after their first Sharpe portfolio is formed.

Once turnover constraints are met for the Sharpe portfolios that fail step 9 but pass what maybe called Step 9a, Steps 10 and 11 are now executed.

This ends the second MOEA.

III. Results

TABLE I
ANNUALIZED RETURNS FOR THE LARGE-CAP MOEA PORTFOLIOS AND THE S&P 500. THE PERIOD COVERED IS FROM DECEMBER 1979 THROUGH DECEMBER 2009 (121 MONTHS).

S&P 500 and MOEA	1 Year	3 Year	5 Year	10 Year
S&P 500	31%	54%	70%	98%
MOEA	53%	92%	119%	168%

TABLE II
RISK AND CUMULATIVE RETURN ON 10,000 USD FOR THE S&P 500 AND THE MOEA PORTFOLIOS. THE PERIOD COVERED IS FROM DECEMBER 1979 THROUGH DECEMBER 2009 (121 MONTHS).

S&P 500 and MOEA	Sharpe Ratio	Information Ratio	Cumulative Return on 10,000 USD
S&P 500	1.1	N/A	10,858
MOEA	1.8	0.14	18,544

TABLE III
ANNUALIZED RETURNS FOR THE LARGE-CAP GROWTH MOEA PORTFOLIOS AND THE RUSSELL 3000 GROWTH. THE PERIOD COVERED IS FROM DECEMBER 1996 THROUGH DECEMBER 2009 (53 MONTHS).

R3000 Growth and MOEA	1 Year	3 Year	5 Year	10 Year
R3000 Growth	4%	8%	10%	14%
MOEA	7%	12%	19%	27%

In terms of percent return and USD return, the MOEA portfolios have double the value of their benchmarks.

On a risk-adjusted basis, the results are not as clear. The information ratio for the large-cap growth MOEA is very

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE IV
RISK AND CUMULATIVE RETURN ON 10,000 USD FOR THE RUSSELL 3000 GROWTH AND THE MOEA PORTFOLIOS. THE PERIOD COVERED IS FROM DECEMBER 1996 THROUGH DECEMBER 2009 (53 MONTHS).

R3000 Growth and MOEA	Sharpe Ratio	Information Ratio	Cumulative Return on 10,000 USD
R3000 Growth	0.14	N/A	18,001
MOEA	0.24	0.32	27,590

significant while its Sharpe ratio is only a little larger than the Russell 3000 Growth Sharpe ratio and both Sharpe ratios are little different from 0 (zero). For the large-cap MOEA, its Sharpe ratio is significantly larger than its benchmark, but its information ratio is very small. So it is not crystal clear that the MOEA portfolios are better, meaner higher, than risk-adjusted returns of the benchmarks. However, by not underperforming their benchmarks on a risk-adjusted basis, the implication is that the MOEAs are at least equal to their benchmarks on a risk-adjusted basis and based upon the risk-adjusted measure used, better than the benchmark. Any portfolio manager would be happy with these results as they have significantly higher performance than their benchmark. And when it comes to risk-adjusted performance versus the benchmark, they are at least the same.

As to the MOEAs ability to meet constraints: 1) In 100% of all cases, the MOEA portfolios on a weighted market capitalization basis meet or exceed the large-cap market capitalization constraint, 2) For the smallest and largest positions constraints based on net asset value, none of the MOEA portfolios broke this constraint on either the minimum or maximum side, 3) Turnover did occasionally exceed the 8% limit per month; these occurrences tended to happen early in the 1980's portfolios - just as the MOEA was getting on its feet - but the turnover limit was also broken in later portfolios as well (The authors surmise that there were times when the non-dominated provided to the second MOEA were composed of stocks different enough from the prior quarter's portfolio that the four-way trade-off between mean, variance, turnover and position limits meant turnover was busted in certain cases. We believe the primary cause for this happening is a significant change in some of the individual stock scores generated by multi-factor model from quarter to quarter. These sorts of changes are very common when using multi-factor models).

IV. CONCLUSIONS

We successfully demonstrate that MOEAs with real world constraints can generate portfolios that exceed the returns of their benchmarks and, at a minimum, equal their benchmarks on a risk-adjusted basis.

We generate these outperforming portfolios by the use of two MOEAs: the first MOEA generates all the non-dominated feasible sets that meet all the constraints except turnover and minimum and maximum position. The second MOEA trades off the last two constraints along with mean and variance to come up with the final portfolio that has the best Sharpe ratio (or meets the maximum turnover allowed if the other constraints are not met).

As best as we know this is the first multi-period use of MOEAs in stock portfolio construction. We are encouraged by the results and hope others can extend and improve upon our work. The authors are now working on extending the unique turnover solution used in the second MOEA to enhance computer security and hopefully improve existing solutions to problems involving worms and viruses. We are doing so by extending the current techniques using homotopy to search for feasible sets. In addition, we are finding faster feasible set results can be computed using linear piecewise topology.

ACKNOWLEDGMENT

The authors would like to thank John Guerard for inviting us to compete in the annual Horse Race sponsored by McKinley Capital. We would not have tested MOEAs in the way described above without John's challenge.

REFERENCES

[1] J. B. Guerard, Jr., Ed., *Handbook of Portfolio Construction*. New York, NY: Springer, 2010, ch. 3.
[2] W. N. Goetzmann, "An introduction to investment theory," http://viking.som.yale.edu/will/finman540/classnotes/notes.html.
[3] Z. Bodie, A. Kane, and A. J. Marcus, *Investments*. Homewood, IL: Richard D. Irwin, Inc., 1989, p. 203 (formula 7.10 and 7.11).
[4] F. Neumann, J. Reichel, and M. Skutella, "Computing minimum cuts by randomized search heuristics," in *GECCO*, C. Ryan and M. Keijzer, Eds. ACM, 2008, pp. 779–786.

ISBN: 978-988-19251-7-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)