Cornell University

Project 2 for ORIE 5370

Higher Order Moments Based Portfolio Selection By Polynomial Goal Programming

Authors:
KaiFeng Gu (kg484)
Bowen Gao (bg453)
Yiheng Xiao (yx426)

Supervisor:
Prof. James Renegar

May 13, 2018

Abstract

The goal of this project is to incorporate realized higher moments into Markowitz Portfolio theory by Polynomial Goal Programming (PGP). First, we used 5-minute returns of components of selected stocks in S&P 600 Small-cap to obtain a daily measure of covariance, coskewnesss and cokurtosis matrices, to analyze the relationship between intraday price movement and future stock returns. Second, under the mean-variance-skewness-kurtosis framework, this study tried to solve multiple conflicting and competing portfolio objectives such as maximizing expected return and skewness and minimizing risk and kurtosis simultaneously. To achieve this goal, we utilized a PGP model which combined four optimization tasks into a single problem, in which investor preferences over higher return moments are incorporated as well. We tested our model on high frequency stock prices dataset and from the results, the Sharpe Ratio of our portfolio outperforms traditional Markowitz Portfolio by 5.37%.

1 Introduction

1.1 Review of Project 1

In project 1, we incorporated realized intraday volatility and skewness as our investor's view into portfolio construction by using Black-Litterman Model to improve the estimation of expected return μ. We took intraday higher moments as trading signals in creating weekly portfolio on DJI 30. Although our Black-Litterman portfolio managed to win over Markowitz, it failed to beat the other benchmark portfolio constructed by equally weighted stocks.

- **Asset Pool**
  - **Weakness** In project 1, we restricted our attention to DJIA 30 component stocks only. We conducted cross-sectional regression and computed information coefficient on higher order moments and future stock returns. However, these statistical tests showed non-significant results. The optimization model did not yield very ideal results on those assets as well. One of the reasons could be that DJI consists of mature, huge companies with high capitalization, which may not be susceptible to anomalies.
  - **Remedy** We shift our focus to small-cap equities (components of S&P 600 Small Cap Index) to see whether higher order moments can be recognized as anomalies on these stocks. Please read section 3.1.1
- **Correlations of skewness and kurtosis between assets**

- **Weakness** In project 1, we merely considered the higher moments of each asset individually without taking the crossing terms of skewness [1] and kurtosis [2] into account. While this kind of information can be crucial for analyzing the co-movement of assets.
- **Remedy** Add co-skewness and co-kurtosis matrices to optimization problem. Details are shown in section 2.2

We start with the traditional Markowitz theory as the building block of our reformed project and one of the benchmarks to beat.

1.2 Improve Protfolio Selection

Markowitz (1952) introduced the mean-variance model (MVM) which has been accepted as the first modern portfolio optimization model. In order to form a strong basis for MVM, there is an assumption with regards to returns of assets that they follow normal distribution. In the literature, the number of studies which cover higher moments in portfolio selection are increasing. [1]

Samuelson [2] also showed that the higher moment is relevant to the investor's decision in portfolio selection. Furthermore, the kurtosis can reflect the probability of extreme events. The larger the kurtosis is, the higher the probability of extreme events happens. It is, thus, extremely necessary to construct portfolio selection within the mean-variance-skewness-kurtosis framework.

Thus, we assume that investors prefer higher expected return, higher skewness, but try to avoid high level variance and kurtosis. This is the core assumption in our portfolio construction and optimization.

Entropy, often taken as a measure of "disorder", can also be used as an objective function in portfolio optimization to ensure diversification. Therefore, we incorporate entropy to this higher order moment portfolio framework to make it a mean-variance-skewness-kurtosis entropy model (MVSKEM).

How to implement portfolio selection with four objectives (four moments) at the same? We resort to a different optimization model called Polynomial Goal Programming (PGP), which will be introduced in the following chapters.

Thus, we divide our optimization programming into two parts.

Multi-objective Programming: We optimize only one objective function at time. To be more specific, we maximize expected return, skewness, and minimize variance and kurtosis

[1] namely, the Co-skewness, $\mathbb{E}[(R_i - m_i)(R_j - m_j)(R_k - m_k)]$

[2] namely, the Co-kurtosis, $\mathbb{E}[(R_i - m_i)(R_j - m_j)(R_k - m_k)(R_l - m_l)]$

respectively in every trading period.

Portfolio Optimization: The aspired value obtained from previous part are substituted into a compound optimization problem with a polynomial objective function. The solution constructs a multi-objective optimal portfolio.

1.3 Structure of the Paper

Section 2 formulates a PGP approach for portfolio selection within the mean-variance-skewness-kurtosis framework.

Section 3 provides an empirical analysis of PGP approach.

Section 4 introduces optimization algorithms to solve PGP.

Section 5 shows all experiment results.

Section 6 concludes the report.

2 Methodology

In this section, we present a multiobjective approach — polynomial goal programming (PGP) — to solve portfolio selection problem within mean-variance-skewness-kurtosis framework.

PGP was first introduced by Tayi and Leonard to facilitate bank balance sheet management with competing and conflicting objectives.[3] Our contribution is to optimize trading strategies within the mean-variance-skewness-kurtosis-entropy framework in constructing portfolio of small-cap equities.

For simplicity, here we assume that R is the distribution of returns and M is expectation of the return, $X = (x_1, x_2, \dots, x_n)^T$ is the weight vector used to combine the portfolio, x_i is the percentage of wealth invested in the i-th risky asset (in our case, investment strategy).

2.1 Important Notations

- **Expected Returns Vector** M

$$M := (m_1, \cdots, m_n)^T, \text{ where, } m_i = \mathbb{E}(R_i)$$

- **Variance-covariance Matrix** V

$$V := \left(\sigma_{ij}\right)_{i,j} = \mathbb{E}[(R_i - m_i)(R_j - m_j)]_{i,j}$$

- **Skewness-coskewness Matrix $S^{\textbf{left}}$ and $S^{\textbf{right}}$**

$$\begin{aligned}S^{\textbf{left}} &:= \left(s_{iij}\right)_{i,j} = \mathbb{E}[(R_i - m_i)^2(R_j - m_j)]_{i,j},\ i,j = 1,\cdots,n\\ S^{\textbf{right}} &:= \left(s_{ijj}\right)_{i,j} = \mathbb{E}[(R_i - m_i)(R_j - m_j)^2]_{i,j},\ i,j = 1,\cdots,n\\ &= (S^{\textbf{left}})^T\end{aligned}$$

- **Kurtosis-cokurtosis Matrix $K^{\textbf{left}}$, $K^{\textbf{middle}}$ and $K^{\textbf{right}}$**

$$\begin{aligned}K^{\textbf{left}} &:= \left(k_{iiij}\right)_{i,j} = \mathbb{E}[(R_i - m_i)^3(R_j - m_j)]_{i,j},\ i,j = 1,\cdots,n\\ K^{\textbf{middle}} &:= \left(k_{iijj}\right)_{i,j} = \mathbb{E}[(R_i - m_i)^2(R_j - m_j)^2]_{i,j},\ i,j = 1,\cdots,n\\ K^{\textbf{right}} &:= \left(k_{ijjj}\right)_{i,j} = \mathbb{E}[(R_i - m_i)(R_j - m_j)^3]_{i,j},\ i,j = 1,\cdots,n\\ &= (K^{\text{left}})^T\end{aligned}$$

In portfolio selection, we focus on the relationship between two assets. Hence, we make a reasonable simplification by defining the higher moments as above instead of the coskewness among three different assets ($s_{i,j,l},\ i \neq j \neq l$) and cokurtosis among three or four different assets.

2.2 Higher Moments of Portfolio

Mean, variance, the third and fourth moments of the return portfolio $X = (x_1, \cdots, x_n)^T$ and entropy measures of portfolio are defined respectively as follows[3]:

- **Expected Return:** $R(x) = X^T M$

- **Variance:** $V(X) = X^T V X$

- **Coskewness:**

$$S(X) = \mathbb{E}[X^T(R - M)]^3 = \sum_{i=1}^{n} x_i^3 s_{iii}^3 + 3\sum_{i=1}^{n}\left(\sum_{j=1}^{n} x_i^2 x_j s_{iij} + \sum_{j=1}^{n} x_i x_j^2 s_{ijj}\right),\ i \neq j$$

- **Cokurtosis:**

$$\begin{aligned}K(X) &= \mathbb{E}[X^T(R - M)]^4\\ &= \sum_{i=1}^{n} x_i^4 k_{iiii}^3 + 4\sum_{i=1}^{n}\left(\sum_{j=1}^{n} x_i^3 x_j s_{iiij} + \sum_{j=1}^{n} x_i x_j^3 s_{ijjj}\right) + 6\sum_{i=1}^{n}\sum_{j=1}^{n} x_i^2 x_j^2 k_{iijj},\ i \neq j\end{aligned}$$

[3] $R = (r_1, \cdots, r_n)^T$ is the observed return of assets, $R(X)$ is a function of portfolio X, representing the expected return.

- **Gini-Simpson Entropy:** $E_s(X) = 1 - X^T X$

Gini-Simpson Entropy is a diversity index which reflects how many different types (such as species) there are in a dataset (a community), and simultaneously takes into account how evenly the basic entities (such as individuals) are distributed among those types.

To enhance diversification in portfolio selection, we use Gini-Simpson Entropy as one of our objectives.

2.3 Sub-problems of PGP

Our proposed model incorporates entropy measures into higher moment portfolio models to provide better dIversification in portfolios. The portfolio objectives can be represented in the multi-objective optimization process (P1) in the following order:

Let $\mathbf{1} = (1, \cdots, 1)^T$.

$$P1 \begin{cases} \max\limits_X \ R(X) \\ \min\limits_X \ V(X) \\ \max\limits_X \ S(X) \\ \min\limits_X \ K(X) \\ \max\limits_X \ E_s(X) := 1 - X^T X \\ s.t. \ X^T \mathbf{1} = 1, \ X \geq 0 \end{cases} \tag{1}$$

Here, $E_s(X) := 1 - X^T X$ is the Gini-Simpson entropy measure. By ensuring shrinkage towards maximum diversification, entropy may lead to better out-of-sample performance. Hence, we add it as one of our objectives.

A simple but effective way to solve $P1$ is to consolidate the various objectives into a single objective function. Hence we use a polynomial goal programming approach to combine these objectives. Let d_1, d_2, d_3, d_4, d_5 be the goal variables which account for the deviations of expected return, variance,skewness and kurtosis from the aspired levels, $R^*, V^*, S^*, K^*, E_s^*$, respectively. The aspired level indicates the best case scenario for a particular objective without considering other objectives. Hence, the aspired levels, $R^*, V^*, S^*, K^*, E_s^*$ can be determined by solving four (independent) subproblems ($SP1, SP2, SP3, SP4, SP5$).

$$SP1\begin{cases}\max\limits_{X} R^*(X) = X^TM \\ s.t.\ X^T\mathbf{1} = 1 \\ \quad X \geq 0\end{cases} \tag{2}$$

$$SP2\begin{cases}\min\limits_{X} V^*(X) = X^TVX \\ s.t.\ X^T\mathbf{1} = 1 \\ \quad X \geq 0\end{cases} \tag{3}$$

$$SP3\begin{cases}\max\limits_{X} S^*(X) = \mathbb{E}[X^T(R-M)]^3 \\ s.t.\ X^T\mathbf{1} = 1 \\ \quad X \geq 0\end{cases} \tag{4}$$

$$SP4\begin{cases}\min\limits_{X} K^*(X) = \mathbb{E}[X^T(R-M)]^4 \\ s.t.\ X^T\mathbf{1} = 1 \\ \quad X \geq 0\end{cases} \tag{5}$$

$$SP5\begin{cases}\max\limits_{X} E_s^*(X) = 1 - X^TX \\ s.t.\ X^T\mathbf{1} = 1 \\ \quad X \geq 0\end{cases} \tag{6}$$

2.4 PGP

The objective of the problems can thus be defined as the minimization of deviations from the ideal scenario set by aspired levels. The general Minkovski distance is a metric often used in finance and economics, for the specification of the objective function in PGP. The computational form of the Minkovski distance is

$$Z = \left(\sum_{k=1}^{m}\left|\frac{d_k}{Z_k}\right|^p\right)^{\frac{1}{p}} \tag{7}$$

where Z_k is the basis for normalizing the k-th goal variable. The value of Z_k is set to the corresponding aspired level in this study. In order to allow the investor exhibiting asymmetric preferences towards the mean, variance, skewness and kurtosis of return, we modify Equation (7) and introduce parameters $\lambda_1, \cdots, \lambda_5$ to indicate the relative preference for mean, variance, skewness, kurtosis and entropy, respectively.

Therefore, we can re-formulate $P1$ to include the goal variables and our modified objective function. Given the investor's relative preference over the mean, variance, skewness and kurtosis of return, the PGP model (P2) can be represented as

$$P2 \begin{cases} \min\limits_{X} & Z = \left|1 + \frac{d_1}{R^*}\right|^{\lambda_1} + \left|1 + \frac{d_2}{V^*}\right|^{\lambda_2} + \left|1 + \frac{d_3}{S^*}\right|^{\lambda_3} + \left|1 + \frac{d_4}{K^*}\right|^{\lambda_4} + \left|1 + \frac{d_5}{E_s^*}\right|^{\lambda_5} \\ s.t & X^T M + d_1 = R^* \\ & X^T V X - d_2 = V^* \\ & \mathbb{E}[X^T(R-M)]^3 + d_3 = S^* \\ & \mathbb{E}[X^T(R-M)]^4 - d_4 = K^* \\ & 1 - X^T X + d_5 = E_s^* \\ & X^T \mathbf{1} = 1 \\ & X \geq 0 \end{cases} \tag{8}$$

We added 1 to all normalized goals to ensure that they are larger than 1 and so that the normalized deviation from the goal which is used as penalty, increases strictly with the exponent.

Solving the PGP problem involves a **two-step** procedure.

1. The aspired levels of $R^*, V^*, S^*, K^*, E_s^*$ for the expected return, variance, skewness and kurtosis, respectively, are obtained from SP1-SP5.

2. These aspired values are substituted into P2, and the minimum value of Z can be found for a given set of investor preferences $\{\lambda_i\},\ (i = 1, 2, 3, 4, 5)$.

The solutions of the subporblems (SP1-SP5) are at least as good as the solution attained by P2, where all objectives R, V, S, K, E_s are considered simultaneously. The values of the goal variables d_1, d_2, d_3, d_4, d_5 are always non-negative. In other words, the goal variables denote the amount of underachievement with respect to the best scenario. The weight set X thus forms the basis for constructing a multi-objective optimal portfolio.

2.5 Explanation of λ_i

The preference parameters $\{\lambda_i\}$ have an explicit economic interpretation in that they are directly associated with the marginal rate of substitution (MRS), where $|Y_j^*|$ [4] is the corresponding denominator in equation (P2).

[4] Y_j^* can be $R^*, V^*, S^*, K^*, E_s^*$ respectively.

$$MRS_{ij} := \frac{\partial Z}{\partial d_i} \bigg/ \frac{\partial Z}{\partial d_j} = \frac{\lambda_i}{\lambda_j} \cdot \frac{|d_i|^{\lambda_i - 1}/|Y_i^*|^{\lambda_i}}{|d_j|^{\lambda_j - 1}/|Y_j^*|^{\lambda_j}}$$

which measures the desirability of foregoing objective $SP(i)$ in order to gain from objective $SP(j)$.

3 Data Processing

We first describe the data collection and cleaning. We then show how the realized higher moments are computed.

3.1 Data

We analyze S&P 600 small cap index from March 5^{th}, 2013 to February 28^{th}, 2018. We choose the 30 stocks with the highest cap in that index.

To compute the daily higher moments, we resort to 5-minute price data.

We collect prices every 5 minutes in each trading day from Bloomberg terminals, and construct five-minute returns for a total of 1257 trading days.

3.1.1 Introduction of Selected Stocks

To improve project 1, in this project, the market cap of selected stocks range from 2.7 to 5.2 billion. We believe these small companies are susceptible to anomalies, say skewness and kurtosis.

3.1.2 Deal with Missing Data

The advantage of analyzing components of S&P 600 small cap index is the completeness of data. Although several tick price data might be missing, it cast no darkness on computing intraday higher moments. Only if the every tick data of a stock are missing in a trading day, we will drop the day out of our consideration.

3.2 Higher Moments

We restructure the data, arrange intraday return of selected stocks by day. By equations in section 2.1, we compute the intraday covariance, coskewness and cokurtosis matrices. We average daily higher moments of every 5 days to compute the weekly ones, which can be used for constructing weekly portfolio.

Table 1: Summary of Selected Stocks

Stock	Sector	Market Cap($K)	Stock	Sector	Market Cap($K)
BCPC	Basic Industries	3,015,771	GDOT	Finance	3,718,487
KS	Basic Industries	3,360,065	SIGI	Finance	3,292,938
TREX	Basic Industries	3,360,612	VAC	Finance	3,145,091
B	Capital Goods	3,076,515	CHE	Health Care	**5,202,976**
IART	Capital Goods	5,085,959	HAE	Health Care	4,591,890
HI	Consumer Durables	2,916,092	NEOG	Health Care	3,713,244
FCFS	Consumer Services	4,046,922	STMP	Miscellaneous	4,451,032
FIVE	Consumer Services	4,030,067	AVA	Public Utilities	3,446,362
PSB	Consumer Services	3,252,173	ASGN	Technology	4,430,730
WWE	Consumer Services	3,323,622	CACI	Technology	4,001,400
PDCE	Energy	3,958,914	JBT	Technology	**2,796,294**
CBU	Finance	3,056,636	KFY	Technology	3,143,651
COLB	Finance	3,048,981	NSP	Technology	3,648,568
FFIN	Finance	3,546,669	QLYS	Technology	3,063,828
GBCI	Finance	3,267,157	SKYW	Transportation	2,933,018

4 Optimization Methods

The objective function in sub-problem (SP3) is a cubic function involving cosknewness matrix. Since skewness is not convex, CVX package fails to solve it. We thus use a traditional optimization method, Newton's Method, to deal with this problem.

4.1 Newton's Method

The following briefly recall Newton's method for solving systems of nonlinear equations without constraints. Define objective function $G(x) := \begin{bmatrix} g_1(x) \\ \vdots \\ g_n(x) \end{bmatrix}$, and its differential $DG(x) := \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_n(x)^T \end{bmatrix}$.

In order to $\min_{\vec{x}} G(x)$, then, the program will approach the optimum iterating as equation (9) with starting point $x^{(0)}$.

$$x^{(k+1)} = x^{(k)} - DG(x^{(k)})^{-1}G(x^{(k)}), \quad k = 0, 1, \cdots, \text{until } ||x^{(k+1)} - x^{(k)}|| < \varepsilon \tag{9}$$

Theorem 1 (Convergence of Newton's Method)**.** *Let x^* be a solution to $G(x) = \vec{0}$. If $DG(x)$*

*is inevitable and the initial iterate $x^{(1)}$ is **sufficiently close to** x^*.*

Then, Newton sequence converges to x^*, and the rate of convergence is quadratic.

4.2 Sequential Quadratic Programming (SQP)

In order to tackle PGP, a constrained nonlinear optimization, by reviewing the literature, we decide to use Sequential Quadratic Programming (SQP), which is a direct ramification of Newton's Method. SQP method[5] solves a sequence of optimization subproblems, each of which optimizes a quadratic model of the objective subject to a linearization of the constraints. If the problem is unconstrained, then the method reduces to Newton's method. If the problem has only equality constraints, then the method is equivalent to applying Newton's method to the first-order optimal conditions, or Karush–Kuhn–Tucker (KKT) conditions, of the problem.

Consider a nonlinear problem of the form (10), with Lagrangian [5]

$$\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) := g(x) - \vec{\lambda}^T b(x) - \vec{\mu}^T c(x).$$

$$\begin{aligned} \min_{\vec{x}} \; & G(x) \\ s.t \; & b(x) \geq 0 \\ & c(x) = 0 \end{aligned} \tag{10}$$

A basic sequential quadratic programming algorithm defines an appropriate search direction d_k as a solution to the quadratic programming subproblem,

$$\begin{aligned} \min_{d} \; & g(x^{(k)}) + \nabla g(x^{(k)})^T d + \frac{1}{2} d^T \nabla^2_{xx} \mathcal{L}(x^{(k)}, \lambda^{(k)}), \mu^{(k)}) d \\ s.t \; & b(x^{(k)}) + \nabla b(x^{(k)})^T d \geq 0 \\ & c(x^{(k)}) + \nabla c(x^{(k)})^T d = 0 \end{aligned} \tag{11}$$

4.3 Deficiency of Newton's method and its followers

While the algorithm converges swiftly, it depends heavily on the initial position of iterations. In practice, there is in general no way to determine whether the condition in Theorem 1 is fulfilled.

Note:

To solve a non-convex programming problem with both equality and inequality constraints, feasible methods are limited. Fortunately, skewness (the third moment) is a twice differentiable

[5] λ, μ are vectors of multipliers.

objective. Fortunately, SQP is still valid here. In our implementation, we set different initial iteration positions, then, to check the if they converge to a same optimum.

5 Experiment Results

5.1 Understanding Investors' Preference Parameters λ_i

Investors' preference is an important factor, which may change under different market scenarios. Moreover, investors' preferences affect the investment strategy. In order to analyze the effect of investors' preferences on portfolio selection, different levels of preferences are tested.

Recall that we have discussed the economic meaning of $\lambda_i,\ i = 1, \cdots, 5$ in section 2.5. λ_i measures the desirability of foregoing objective $SP(i)$ in order to gain from objective $SP(j)$. In PGP Optimization problem $P2, \min_X Z = \left|1+\frac{d_1}{R^*}\right|^{\lambda_1} + \left|1+\frac{d_2}{V^*}\right|^{\lambda_2} + \left|1+\frac{d_3}{S^*}\right|^{\lambda_3} + \left|1+\frac{d_4}{K^*}\right|^{\lambda_4} + \left|1+\frac{d_5}{E_s^*}\right|^{\lambda_5}$, if λ_i is bigger, then, the distance between Optimized portfolio will be much more closer the respective aspired target $R^*, V^*, S^*, K^*, E_s^*$.

To verify this effect, first, we set $(\lambda_1, \lambda_2, \cdots, \lambda_5) = (1, 1, 0, 0, 0)$ as the benchmark case, which is an equivalent form of Markowitz mean-variance portfolio. We present the results in table 2.

We base our experiment on `'2013-05-01'`. The first column reports parameters λ_i, Second column, the current higher moments of Optimal portfolio optimized based on information on `'2013-05-01'`. Then, with the same strategy, we trade assets for consecutive 10 weeks [6].

The more importance investors' preferences attach to a certain moment, i.e., the greater the preference parameter for this moment, the more favorable value of this moment statistic would be in the optimal portfolio. That is, as a result of the tradeoff between the four moments, at least one of the other three moment statistics deteriorates.

Consider portfolio A, B, C, E, G, I, J, we check the individual effect of each paramter λ_i. For instance, the lowest variance is achieved in Portfolio G. It has higher aversion[7] over variance ($\lambda_2 = 3$). The highest skewness is achieved in Portfolio I. It has higher preference over skewness ($\lambda_3 = 3$).

Consider portfolio A, C, we know that if we add entropy term to PGP, the portfolio can reach a higher return and Sharpe Ratio. It will be verified again in section.

[6] Though we can extend the conclusion for a longer period, the computation is time-consuming. We will conduct the best strategy for entire 249 weeks, which will surely show a detailed performance.

[7] We set $\min x^T V x$ as our objective in SP(2)

Table 2: Asset allocations and moment statistics for optimal portfolio with different preferences

Portfolio	A	B	C	D	E
λ_1	1	1	1	1	3
λ_2	1	1	1	0	1
λ_3	1	0	1	1	1
λ_4	1	0	1	1	1
λ_5	1	0	0	1	1
Mean	0.00651	0.08378	0.00653	0.00651	0.00690
Variance ($\times 10^{-7}$)	3.14448	151.14060	3.14325	3.14998	3.13214
Skewness	-0.00458	-0.05071	-0.00449	-0.00452	-0.00497
Kurtosis	0.01135	1.88843	0.01135	0.01135	0.01144
Annual Return	0.90489	0.10688	0.90304	0.90566	0.88800
Sharpe Ratio	0.62545	0.25492	0.62449	0.62565	0.60477
Portfolio	F	G	H	I	J
λ_1	3	1	1	1	1
λ_2	1	3	3	1	1
λ_3	3	1	1	3	1
λ_4	1	1	3	1	3
λ_5	1	1	1	1	1
Mean	0.00689	0.00653	0.00936	0.00634	0.00938
Variance ($\times 10^{-7}$)	3.16944	3.14275	3.46785	3.16746	3.44616
Skewness	-0.00407	-0.00450	-0.00149	-0.00370	-0.00158
Kurtosis	0.01147	0.01135	0.01347	0.01135	0.01347
Annual Return	0.87662	0.90351	0.77906	0.88737	0.77647
Sharpe Ratio	0.60082	0.62504	0.53886	0.61976	0.53765

5.2 The Change of Moments By Different Preference

If mathematically feasible, optimizing several objectives simultaneously and then solving optimization problem could reduce a lot of work and be much more painless. However, by examining the following examples, we find that the objectives to be optimized in our problem are conflicting with each other. And thus we are not able to solve them at the same time.

We choose May 1^{st}, 2013 as our reference time point. By analyzing stock data in the previous week, we could obtain a optimal trading strategy of this week by PGP algorithm. By assigning different values to $\{\lambda_i\}$, we derived 8 trading strategies. Then, we trade the assets on May 1^{st} and compute the mean, variance, skewness and kurtosis of the portfolio as shown in table 3.

The tradeoff between four moments with different preferences is now articulated in Table 3. In Portfolio b, compared with a, Investors' higher preference for expectation leads to a strategy with not only higher current return but also lower skewness than Portfolio b. Thus, as the investor preference for expected returns increases, he/she must accept a lower skewness. That means we cannot attain higher first and third moment of a portfolio at the same time. They are contradictory goals.

In portfolio d, compared with c , a larger λ_3 leads to a higher portfolio skewness but a lower portfolio kurtosis. Similarly, in portfolio e and f, we change (λ_i) from $(1,1,0,3,0)$ to $(2,1,0,3,0)$ while holding the values of skewness unchanged. This results in the mean rising from 0.00579 to 0.00630 and the kurtosis rising from 0.01124 to 0.01129. Thus, as λ_1 increases, the investors must settle for higher kurtosis, holding skewness constant. When skewness and kurtosis are held constant, as shown in Portfolio g and h, higher preference for expected return leads to higher returns, but also higher variance, which are the same to traditional Markowitz's mean-variance model.

Therefore, expected return and skewness, variance and kurtosis are conflicting pairs of objectives in portfolio diversification for risk averse investors who prefer portfolios with high skewness and expected return but lower variance and kurtosis.

Table 3: The change of moment statistics with the change of preferences

Portfolio	a	b	c	d	e	f	g	h
λ_1	1	2	3	3	1	2	1	2
λ_2	1	1	1	1	1	1	3	3
λ_3	3	3	1	2	0	0	0	0
λ_4	0	0	0	0	3	3	1	1
λ_5	0	0	0	0	0	0	0	0
Mean	-0.00268	0.06054	0.07780	0.06734	0.00579	0.00630	0.00614	0.00744
Variance($\times 10^{-7}$)	69.82633	68.59850	125.52574	88.10590	3.10656	3.10186	3.09933	3.13882
Skewness	0.58354	0.41765	0.15883	0.36642	-0.00538	-0.00538	-0.00539	-0.00521
Kurtosis	2.20235	1.26661	1.82824	1.48813	0.01124	0.01129	0.01127	0.01166

5.3 Performance of PGP Strategy

Note that if we set $\lambda_3, \lambda_4, \lambda_5 = 0$, our PGP model (P2 in equation (8)) is equivalent to Markowitz. With the optimal solutions to the individual objectives $(SP(1), \cdots, SP(5))$, we can solve $P2$ with the PGP approach. For exposition purposes, the optimal sets used to construct the portfolio are reported below.

Given the investors' preference of $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) = (1,1,1,1,1)$ is worth noticing, since, it not only express a indifferent preference on expected return, variance, skewness and kurtosis, but also be considered as a canonical model.

On the first trading day of each week, we summarize the market information (higher moments) in last week. To gauge the historical information comprehensively, we use 5-minute

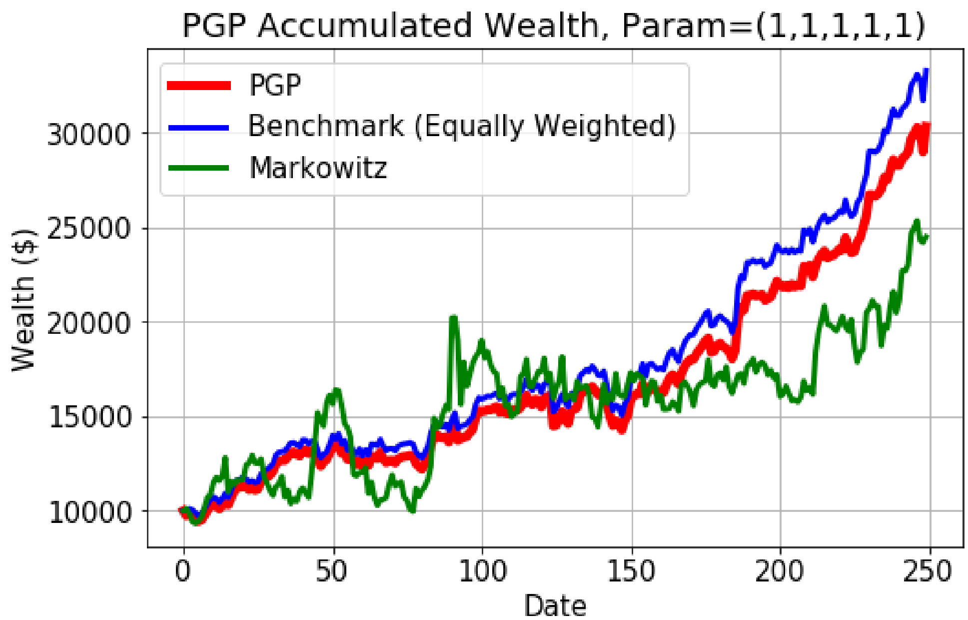

Figure 1: PGP(1)

intraday data. Then, we solve the PGP to decide the new position we should hold in the following week. That means we rebalance every week. Our backtest covers from March 5^{th}, 2013 to February 28^{th}, 2018, totally 1257 trading day, 250 trading weeks.

Other crucial parameters:

- transaction cost = 0.001
- initial wealth = 10000

For comparison, we also report the performance of Markowitz mean-variance strategy and equally weighted portfolio. Table 4 summarizes statistics of strategies. It is obvious that we have defeated Markowitz and approaching the performance of equally weighted strategies.

Table 4: Performance of Strategies

	PGP(1)	PGP(2)	Equally Weighted	Markowitz
Parameters	$(1, 1, 1, 1, 1)$	$(1, 3, 1, 1, 1)$	-	-
Initial Wealth	10000.00	10000.00	10000.00	10000.00
Terminal Wealth	30332.03	30318.04	33297.30	24504.97
Annualized Return	**0.2552**	0.2551	**0.2794**	0.2015
Sharpe Ratio	**0.2205**	0.2204	**0.2368**	0.0927

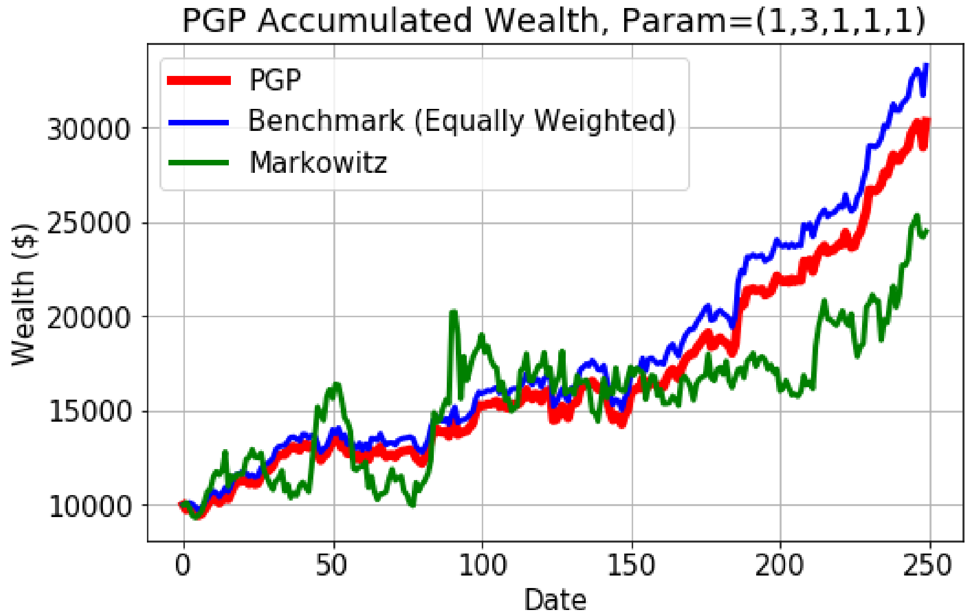

Figure 2: PGP(1)

6 Conclusions

This project has incorporated investor preferences into a PGP optimization function. This approach allows us to solve for multiple conflicting portfolio objectives within the mean-variance-skewness-kurtosis-entropy framework. Our empirical analysis shows the PGP approach can efficiently solve portfolio problem with multiple conflicting objectives and can find optimal portfolio and make the corresponding investment decisions. Meanwhile, empirical results also reveal that the different investors' preferences can affect asset allocations of portfolio or investment strategies.

Finally, we construct PGP strategy with optimal parameters. In the past 5 years, our strategy outperforms the Markowitz and emulates the equally weighted strategy.

References

[1] Arditti, F. D. , Levy, H. (1975). Portfolio efficiency analysis in three moments: The multiperiod case. The Journal of Finance, 30 (3), 797–809.

[2] P. Samuelson, “The fundamental approximation theorem of portfolio analysis in terms of means variances and higher moments,” Review of Economic Studies, 1958, 25, pp. 65-86.

[3] G.K. Tayi and P.A. Leonard, “Bank balance-sheet management: An alternative multi-objective model,” Journal of Operational Research Society, 1988, 39, pp. 401-410.

[4] Frank J. Fabozzi, Petter N. 2007, Robust Portfolio Optimization and Management. John Wiley & Sons, Inc. Chapter 3

[5] Bonnans, J. Frédéric; Gilbert, J. Charles; Lemaréchal, Claude; Sagastizábal, Claudia A. (2006). Numerical optimization: Theoretical and practical aspects

[6] Mehmet Aksaraylı, Osman Pala, A polynomial goal programming model for portfolio optimization based on entropy and higher moments, Expert Systems with Applications,Volume 94, 2018, Pages 185-192

[7] Kin Keung Lai, Lean Yu, and Shouyang Wang, Mean-Variance-Skewness-Kurtosis-based Portfolio Optimization, First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS’06). Volume 2, Pages 292-297.