Linear and Quadratic Portfolio Optimization: A Brief Study

Zane Jakobs, Sarah De Nardis, Michael Wegner

April 29, 2019

1 Introduction

In the world of finance– rather, in the idealized world of finance– the greatest long term gains are obtained by simultaneously minimizing risk and maximizing returns. However, in reality, these ends are usually in diametric opposition to each other: the greatest returns (and, inversely, the greatest losses) require a higher level of risk, whereas, conservative risk-taking yields conservative losses (and, inversely, conservative gains). First, however, we must define risk in a way that we can both measure and control it–at least to some degree.

In general, we define risk with respect to two metrics: portfolio standard deviation, usually measured as a percentage, and market exposure. Portfolio standard deviation can be controlled with Markowitz portfolio optimization (and will be discussed in greater detail later), and, since we are using a long-short equity strategy, market risk can be controlled by limiting something called dollar exposure.

Dollars may be allocated to different market exposure positions, called long and short positions. Long positions refer to the buying of company shares with the expectation that the company's share market prices will increase and shares can be sold at a profit. There's no time limit for how long a long position can be held. The risk associated with such investments, of course, is that the investor stands to lose his investment (though no more than the amount he invested) if the company's stock prices drop. Short dollar positions refer to borrowing shares from another shareholder with the promise to replace those shares within an agreed upon frame of time. The buyer, with the expectation that the market prices drop, will then short-sell these shares, and after the market prices drop, buy back the shares at a lower cost to return to the lender and pocket the profit. Whereas the long position holder increases profit as the price of stock goes up, the short position holder increases profit as the price of stock goes down. The danger inherent with short position holding is that if the price of stock has gone up by the end of the agreed-upon repayment period, the investor incurs a loss equal to the difference in the new share price and the borrowed share price– a loss that is only limited by how high the stock price may go.

2 The Linear Program

2.1 Background: Financial Terminology

Consider a portfolio with a proportion of $0 < f < 1$ available capital (dollars) invested in long positions and $1 - f$ in short positions. Each security has its own market exposure, measured with β_m, defined as the ordinary least-squares correlation coefficient,

$$\beta_m = \frac{\mathrm{Cov}(r_s, r_m)}{\mathrm{Var}(r_m)}. \tag{1}$$

This coefficient can be interpreted as how much the given security's price moves with or against the market per a unit change in the market price. With a sufficiently diversified portfolio (i.e., one that contains many assets with hopefully fairly low intra-portfolio correlation), the average market beta of chosen long positions equals the average market beta of chosen short positions. If the fund manager chooses long positions that return on average $M + a$, where M is the market return, and short positions with average return of $M - b$, then the portfolio's returns are[1]

$$r_p = f(M + a) - (1 - f)(M - b) = (2f - 1)M + fa + (1 - f)b. \tag{2}$$

In the special case of $f = 0.5-$ that is, a portfolio with equal amounts of capital invested in long and short positions, and under the assumption that the portfolio is sufficiently diversified such that the average market beta of long and short positions is the same, equation (2) then reads

$$r_p = \frac{a + b}{2}, \tag{3}$$

and thus is independent of the market's returns.

[1]In mathematical finance, short positions are represented by negative values, as the equity's value going down corresponds to the portfolio's value going up. The reader should ensure this makes sense to him-/herself before continuing.

2.2 Linear Program Implementation

In light of equation (2), we now define the portfolio's market beta as

$$\beta_m = \frac{D_L - D_S}{V_0}, \tag{4}$$

where D_L denotes the total dollars occupying long positions, D_S denotes the total dollars occupying short positions, and V_0 denotes the portfolio's initial value. The larger the value of β_m then, the greater the market risk assumed. Note that β_m may be a positive or a negative value: positive if long dollar positions exceed short dollar positions and negative if short dollar positions exceed long dollar positions.

We thus wish to optimize expected returns under constraint of a maximum market beta, $\beta_{\max}$, and a minimum market beta, $\beta_{\min}$, and both or either may be positive or negative. Our linear program accepts a vector of expected returns, where each return is calculated over a period of time, i, P_i denotes the initial price of the security, and P_f denotes the price of the same security at the end of time period i:

$$r[i] = \frac{P_f - P_i}{P_i}.$$

Note that a positive return reflects an uptick in price, and a negative return reflects a downturn. As our decision variables are then a vector of weights (or percentage investment allocations in each security), a positive weight corresponds to a fraction of $\beta_{\max}$ that, when paired with a positive return, will result in a gain and when paired with a negative return will result in a loss. The situation is reversed with short positions, where negative returns correspond to gains, and positive returns correspond to losses. In general, a positive dollar exposure, and hence positive portfolio market beta, means the portfolio's value will increase when the market increases, and decrease when the market decreases, and a negative portfolio market beta implies the opposite. A market beta of zero implies that the portfolio's returns are entirely independent on the market's returns. In general, real long-short equity funds often run market betas between 0.3 and 0.7 in order to gain from the long term positive expected returns on the market (i.e., we expect the market's returns to be positive over, say, ten years). However, there is also a niche for so-called market neutral funds, which run market betas as close to zero as possible. Funds running a negative market beta are rare, but exist, usually in the context of merger arbitrage funds. We chose to optimize our portfolios with the constraint that our dollar exposure– and hence, our market beta– was between -0.3 and $+0.3$.

Our objective function is further constrained by limiting the sum of the absolute values of weights to equal leverage, where leverage is defined as the sum of total owned capital and total borrowed capital divided by total owned capital. A maximum position size is also applied as a constraint to each weight in order to ensure a sufficiently diverse portfolio.[2]

We now present the linear program specification: given weights (decision variables) $\mathbf{w}$, we first define $\mathbf{w}_+$ as the positive weights in the portfolio (long positions) and $\mathbf{w}_-$ as the negative weights, such that $\mathbf{w} = \mathbf{w}_+ + \mathbf{w}_-$. For example, if the optimal weight of ticker AAPL is 0.02, then $\mathbf{w}_{+,\text{AAPL}} = 0.02$, and $\mathbf{w}_{-,\text{AAPL}} = 0$. This is necessary to linearize the leverage constraint, which states that the absolute values of the weights sum to less than or equal to the maximum permissible leverage. Furthermore, let $\mathbf{r}$ be the vector of expected returns. We then have the

[2] Individual positions were capped at 5% of total portfolio value. A more sophisticated optimization scheme might use the Kelly criterion, or a similar measure of maximum allowable risk. However, to actually implement such a strategy requires higher-frequency price data than we were able to obtain for all equities in the S&P500, on the order of 60-minute returns [2].

linear program

$$
\begin{aligned}
\max\ &\mathbf{r}^T\mathbf{w} \\
\text{s.t.}& \\
\mathbf{1}^T\mathbf{w} &\geq \text{MinDollarExposure} \\
\mathbf{1}^T\mathbf{w} &\leq \text{MaxDollarExposure} \\
\mathbf{1}^T(\mathbf{w}_+ - \mathbf{w}_-) &\leq \text{MaxLeverage} \\
\forall i,\ 0 \leq \mathbf{w}_{+,i} &\leq \text{MaxPosSize} \\
\forall i,\ -\text{MaxPosSize} \leq \mathbf{w}_{-,i} &\leq 0
\end{aligned}
\tag{5}
$$

which, for the equities we considered (those that have been in the S&P 500 since January 4, 2010, and remained in the S&P 500 through at least December 28, 2018), gives 455 decision variables and $2 \cdot 455 + 3 = 913$ constraints. For speed, the Gurobi optimizer was used as our linear solver, and various time series plots where generated for analysis.

3 The Quadratic Program

3.1 Background: Markowitz Mean-Variance Portfolio Theory

The concept of maximizing return in light of market risk is expanded upon by the Markowitz Mean-Variance Model. Whereas our linear program used a defined market beta to account for dollar exposure risk, this model goes a step further to incorporate market volatility in the risk assessment. This market volatility is measured by the standard deviations of expected returns for a set of assets over a specified time period. The objective, then, is to maximize returns, subject to an accepted level of risk, given by a specific standard deviation (being the square root of the return variance).

In general, the Markowitz algorithm accepts as inputs the expected returns for a set of assets and their expected covariance matrix. In order to simulate various forecast accuracy levels, we injected Gaussian noise[3] into the realized returns for the dates with which we optimized.

It is important to use a robust estimator for the covariance matrix as, in general, the distribution of returns is not Normal. We take advantage of the fact that if C is the correlation matrix of an indexed collection of random variables and S is a diagonal matrix whose nonzero entries are the standard deviations of the random variables (in the same order), then the covariance matrix Σ is

$$\Sigma = SCS. \tag{6}$$

Using (5), we can estimate the correlation matrix in a time-stable and distribution-independent manner, as well as the standard deviations, with an estimator that accounts for the observed conditional heteroskedasticity and non-normality of returns distributions.

A better distributional assumption than the Normal would be that of an α-stable distribution with $\alpha \approx 1.8$ [3]. (The Normal is also a stable distribution, but with $\alpha = 2$.) However, we obtained our price data from the Yahoo! Finance API through-wrapper provided by the R package tidyquant, which does not provide sufficiently high-frequency data to accurately fit stable distributions to the relevant equities. In light of these challenges, we modeled the variances of equities (and hence standard deviations) as a generalized asymmetric student-t (AST)

[3] In light of questions asked during the presentation, we also tried using uniformly distributed noise, which did not provide more realistic data. The reason Gaussian noise works is that a Gaussian distribution, while visibly wrong– see figure 4– is a semi-reasonable estimate of the returns distribution and, hence, the distribution of the difference between a point forecast and the observed value (by the Levy-Lindeberg Central Limit Theorem). A uniform distribution, on the other hand, is completely different from the real forecast-minus-observation distribution, which is likely best approximated with an α-stable distribution that isn't the Normal [2]. Additionally, changing the mean of the noise distribution would amount to a change in the forecast of the market's returns, which is irrelevant if market beta is zero. Even if market beta is relatively high, this does not change the optimal portfolio significantly, particularly if market returns are controlled for (e.g., with the Capital Asset Pricing Model (CAPM)).

distribution [4], which was taken from an equities return distribution estimation we implemented using a Generalized Autoregressive Score (GAS) model [1].

We can now present the quadratic program specification of Markowitz portfolio theory. Note that $\mathbf{w}_+$ and $\mathbf{w}_-$ are defined as in the linear program specification, and for the same reason. We then define the *risk tolerance*, q, as a conversion factor between risk (in terms of standard deviation of returns) and returns, which depends on the investor's risk preferences. Letting Σ be the covariance matrix of returns and $\mathbf{r}$ be a vector of expected returns, we have the program specification

$$\begin{aligned} \min\ & \mathbf{w}^T \Sigma \mathbf{w} - q \mathbf{r}^T \mathbf{w} \\ \text{s.t.}\ & \\ \mathbf{1}^T \mathbf{w} &\geq \text{MinDollarExposure} \\ \mathbf{1}^T \mathbf{w} &\leq \text{MaxDollarExposure} \\ \mathbf{1}^T (\mathbf{w}_+ - \mathbf{w}_-) &\leq \text{MaxLeverage} \\ \forall i,\ 0 \leq \mathbf{w}_{+,i} &\leq \text{MaxPosSize} \\ \forall i,\ -\text{MaxPosSize} \leq \mathbf{w}_{-,i} &\leq 0 \end{aligned} \tag{7}$$

which, like the linear program, has 455 decision variables and 913 constraints.

Varying q yields the efficient frontier, which is the set of portfolios with expected return greater than any other with the same or lesser risk, or lesser risk with the same or greater return. A simplified explanation of the efficient frontier is given in the case of allocations to two assets. Say asset 1 has an expected return of 10% and a standard deviation of 10%, and asset 2 has an expected return of 13% and a standard deviation of 30%. Also consider that these assets are 100% negatively correlated, meaning when asset 1 is yields a return above its expected value, asset 2 is 100% likely to yield a return below its expected value (and vice versa). For example, if a standard deviation of zero is desired (meaning zero risk), a 75% allocation of asset 1, coupled with a 25% allocation of asset 2 will yield a standard deviation of zero. That is, $0.75 * 10 = 7.5$ and $0.25 * 13 = 7.5$, the difference of which (due to perfect negative correlation) is zero. The Markowitz Mean-Variance algorithm proceeds in like matter to optimize returns subject to a standard deviation (risk) designation.

3.2 Data and Results

We optimized various portfolios consisting of all 455 considered equities with a minimum dollar exposure of -0.3, a maximum dollar exposure of $+0.3$, a maximum position size of 5% of total portfolio value, and a risk tolerance of 0.5, changing the noise parameter used across each run over a given time interval. The portfolios were initialized with $\$100,000$, and results are shown below.

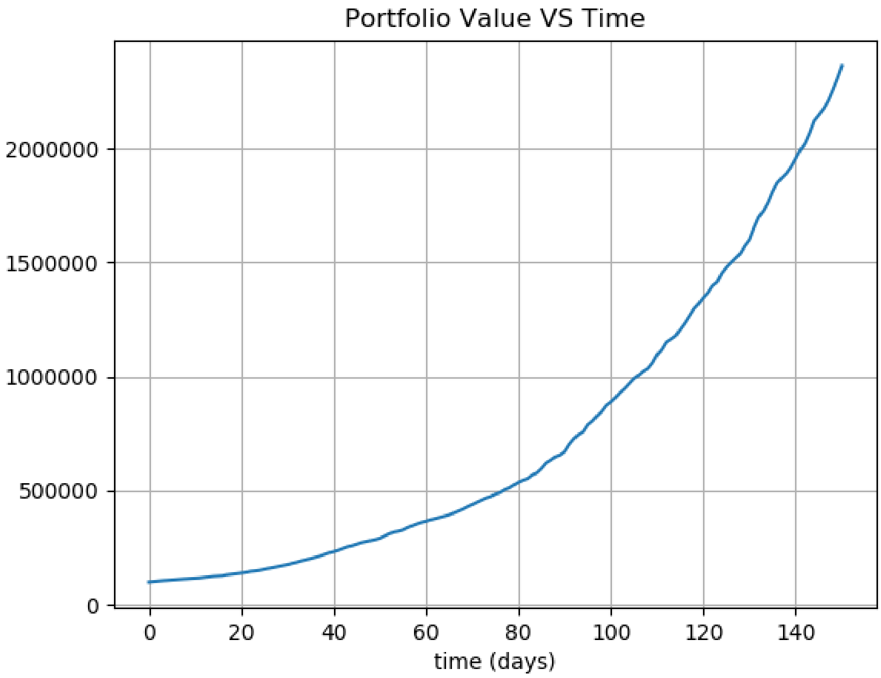

Figure 1: Markowitz-optimal portfolio values over 150 trading days with Gaussian noise and noise parameter 2. The extreme returns shown here made us consider injecting more noise into the future returns.

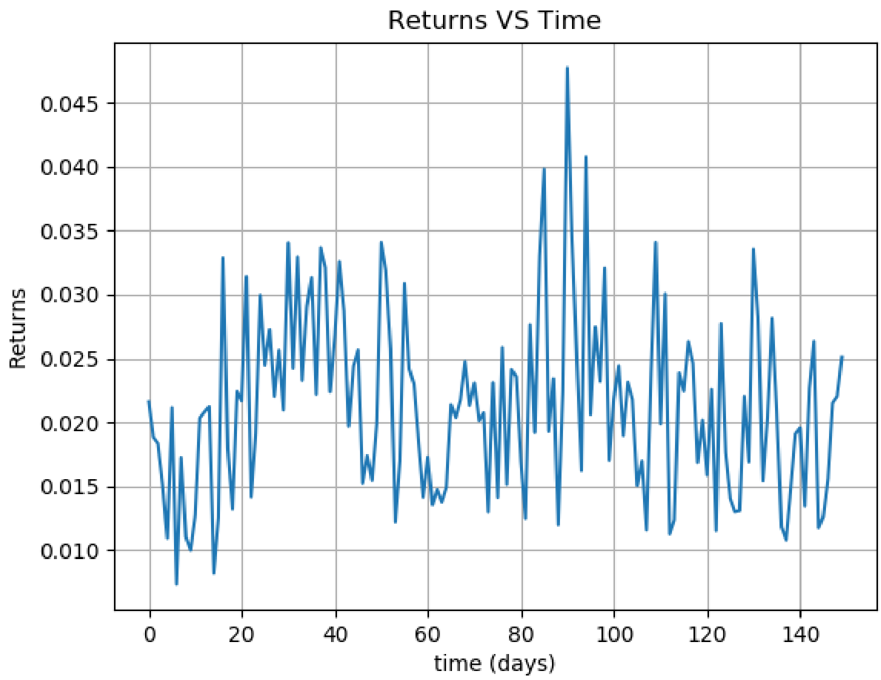

Figure 2: Markowitz-optimal portfolio returns over 150 trading days with Gaussian noise and noise parameter 2. Note that no trading days have negative returns, and the conditional heteroskedasticity of stock returns is reflected in the portfolio return. While the injected noise is i.i.d., it was not large enough to significantly worsen the optimizer's performance to a realistic level. Over short periods, however, volatility is relatively low.

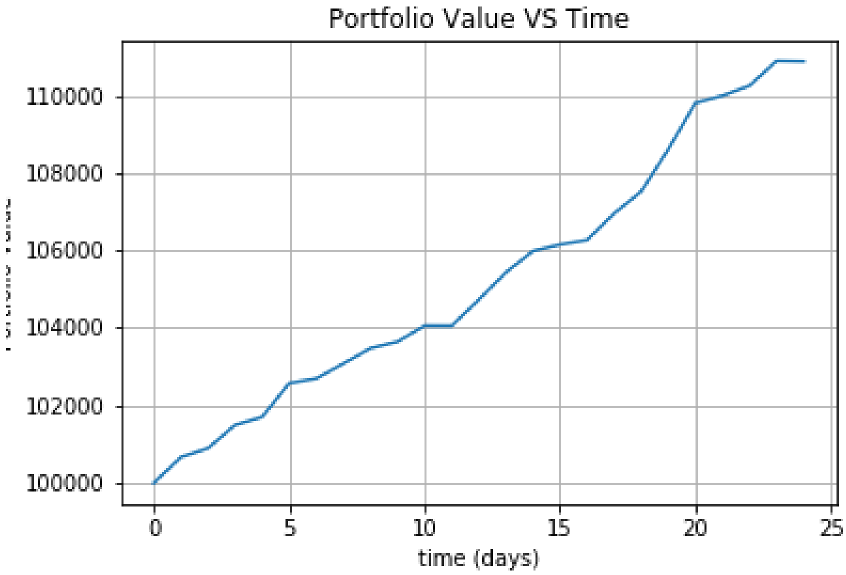

Figure 3: Markowitz-optimal portfolio values over 25 trading days with Gaussian noise and noise parameter 8. These returns, while still high (approximately 11.8%–which many funds would consider a successful annual return–over 25 days), are more realistic than with smaller noise parameters and actually challenge the model to reduce the variance of returns.

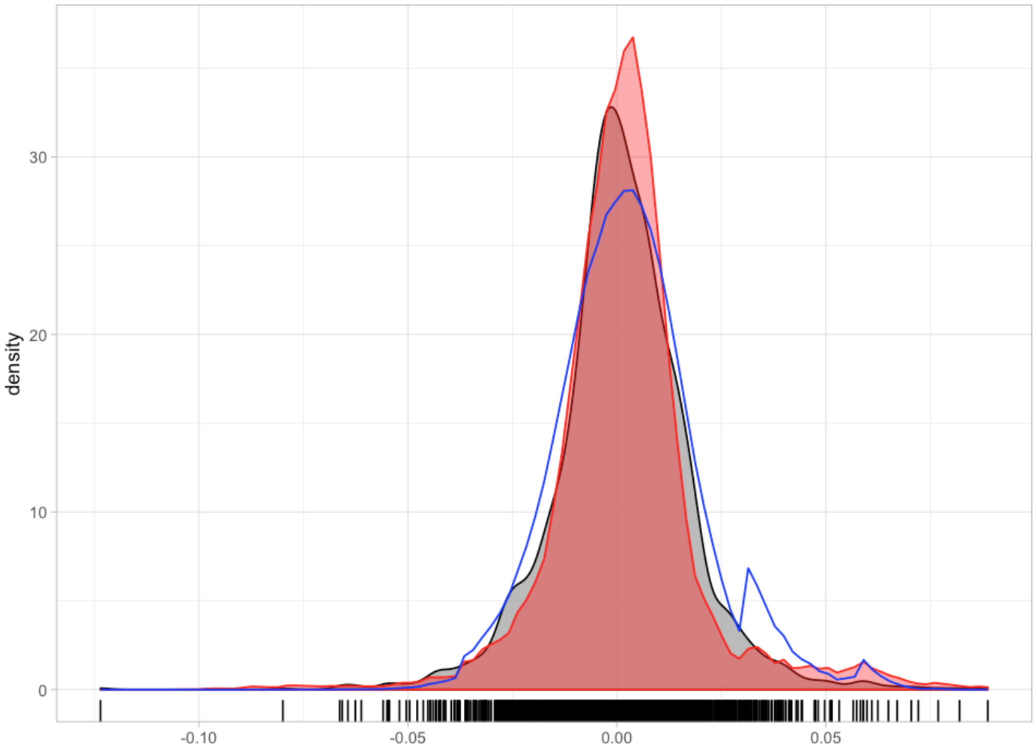

Figure 4: Example of distribution of returns (grey, with individual days represented by black tick marks at bottom) fit with Asymmetric Student-t distribution (red) and Normal distribution (blue). Distributions were fit with a Generalized Autoregressive Score model and reflect the distribution over time (with varying parameters at each time step).

4 Conclusion and Future Work

As we have shown, Markowitz portfolio optimization can produce extremely high-return, low-risk portfolios if the optimizer is given good enough forecasts of the expected return, and accurate estimates of the covariance. Additionally, we have built an extensible Markowitz optimizer with the Gurobi optimization framework, which can do both linear portfolio optimization (by either letting the covariance matrix be zero or setting the risk tolerance to a very large, effectively infinite number) and quadratic optimization in the Markowitz mean-variance framework for long-short equity portfolios with arbitrary leverage and dollar exposure constraints. Additionally, should a portfolio manager want to limit other factor exposures, incorporating such a change would be as uncomplicated as imposing that the sum of squared weights times squared factor betas not exceed (the square of) a desired maximum. Such a constraint–and indeed any quadratic constraint that can be expressed with multiplication by a positive-semi-definite matrix–can be readily handled by the solver. Other possible model extensions include risk modeling with Value-at-Risk (VaR), optimization with respect to minimizing the expected shortfall (the expected value of portfolio losses, given that losses exceed the VaR for a given threshold), optimizing the logarithm of wealth instead of regular returns (which yields a Kelly-optimal portfolio), and considering higher cumulants of the returns distribution than only the mean and variance (e.g. skewness and kurtosis).

References

[1] M. Bernardi and L. Catania. Portfolio optimisation under flexible dynamic dependence modelling. *Journal of Empirical Finance*, 48:1–18, 2018.

[2] M. de Pooter, M. Martens, and D. van Dijk. Predicting the daily covariance matrix for s&p 100 stocks using intraday data - but which frequency to use? *Tinbergen Institute*, 2005. URL `https://papers.tinbergen.nl/05089.pdf`.

[3] A. Reuss, P. Olivares, L. Seco, and R. Zagst. Risk management and portfolio selection using α-stable regime switching models. *Applied Mathematical Sciences*, 10(12):549–582, 2016. doi: http://dx.doi.org/10.12988/ams.2016.512722.

[4] D. Zhu and J. W. Galbraith. A generalized asymmetric student-t distribution with applications to financial econometrics. *CIRANO*, April, 2009. URL `https://pdfs.semanticscholar.org/d91f/bda26e5824717245e35621e961885cbee2b3.pdf`.