Explicit Solution for Position Constrained Markowitz Problems

Alex Bernstein
abernstein@ucsb.edu
joint work with Alexander Shkolnik

SIAM Conference on Financial Mathematics and Engineering
June 4, 2019

Department of Statistics & Applied Probability
University of California, Santa Barbara

The Markowitz Problem

The Markowitz Minimum Variance Portfolio solves:

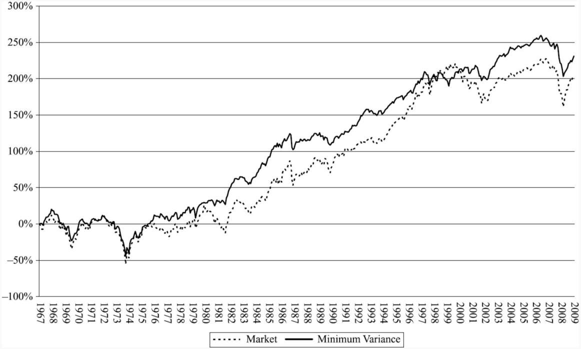
$$\begin{aligned} &\min_{x \in \mathbb{R}^p} x^\top \Sigma x \\ &\text{subject to:} \\ &\left\{ \; \textstyle\sum_k x_k = 1, \right. \end{aligned} \tag{1}$$

where Σ is the covariance matrix for the vector of securities x.
Explicit Solution:

$$x = \frac{\Sigma^{-1}\vec{1}}{\vec{1}^\top \Sigma^{-1} \vec{1}} \tag{2}$$

Performance of Minimum Variance Portfolio

Percent Return Over Time of Market vs. Minimum Variance Portfolio for 1000 Largest US Stocks



(Clarke, de Silva & Thorley 2011)

- Minimum Variance Portfolio has usually beaten Market Portfolio (1000 largest US Equities). This is the Low-Volatility anomaly: low beta stocks get a higher than expected return.

Long-Only Markowitz Problem

The Long-Only Markowitz Minimum Variance Portfolio* solves:

$$\min_{x \in \mathbb{R}^p} x^\top \Sigma x$$

(3) subject to:

$$\begin{cases} \sum_k x_k = 1, \\ \text{every } x_k \geq 0. \end{cases}$$

* We assume no minimum return constraint for notational simplicity where Σ is the covariance matrix for the vector of securities x.

- Long-only constraint makes the model **investable**, as it does not require constant rebalancing.
- No Explicit Solution in general; **requires numerical optimization.**

Motivation

Why look for explicit solutions?

- For $\Sigma = \sigma^2 \beta\beta^\top + \Delta$ where $\beta \in \mathbb{R}^p$ and Δ is a diagonal $p \times p$ matrix of specific variances δ_i^2 (CAPM, i.e. a one-factor model), solution is:

$$x_i = \begin{cases} \frac{\sigma^2}{\delta_i^2}\left(1 - \frac{\beta^{(i)}}{\beta^*}\right) & \text{if } \beta_i < \beta^* \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

  where β^* is a cutoff value computed as a fixed point (Clarke et al. 2011)
- Want a more formal proof and more flexibility as empirical evidence suggests there is more than one factor.

Structural Model

Structured Covariance matrix Σ with M factors:

$$\Sigma = BB^\top + \Delta \tag{5}$$

$$B = \underbrace{\begin{bmatrix} \vdots & \vdots & & \vdots \\ \beta_1 & \beta_2 & \dots & \beta_M \\ \vdots & \vdots & & \vdots \end{bmatrix}}_{p\times M}, \qquad \Delta = \underbrace{\begin{bmatrix} \delta_1^2 & & \\ & \ddots & \\ & & \delta_p^2 \end{bmatrix}}_{p\times p}$$

Model assumes $M << p$.

- Interested in solutions with more factors
- Want to analyze the underlying mathematics of minimum variance factor models

Structural Model

Why structure a Covariance Matrix?

- Factor Analysis/Low-dimensional approximation; can be estimated with statistical methods such as PCA or MLE
- Compatible with Pricing theory such as CAPM and Arbitrage Pricing Theory (APT). For example:
    - CAPM: $M = 1$
    - Fama-French: $M = 3$

Long-Short under Structural Model

Note that:

$$\min_{x \in \mathbb{R}^p} x^\top \Sigma x = \min_{x \in \mathbb{R}^p} x^\top \boldsymbol{B}\boldsymbol{B}^\top x + x^\top \Delta x \tag{6}$$

both are subject to $x^\top \vec{1} = 1$. We have a tradeoff between:

- Hedge Risk Factors in $\boldsymbol{B}$
- Stay fully invested ($x^\top \vec{1} = 1$)

Further, we have that:

$$x \propto \vec{1} - \sum_{j=1}^{k} \theta_j \beta_j \tag{7}$$

Notation

We use the following notation:

- M- number of factors
- p- number of securities
- $\beta_k^{(i)}$- i^{th}element of k^{th}factor
- x_i- i^{th}element of the weight vector, x
- δ_i^2- i^{th}specific variance (specific variance of i^{th}security)
- σ_k^2- variance of k^{th}factor
- $\theta_k = \frac{\sigma_k^2 \left(\sum_{j=1}^{p} \beta_k^{(i)} x_i \right)}{\ell}$ (ℓ is the Lagrange multiplier of our fully invested constraint- not important for the solution)

Important Functions

Define the following functions of θ:

$$a_k(\theta) = \frac{1}{\sigma_k^2} + \sum_{\{i : f_i(x,\theta)<1\}} \frac{(\beta_k^{(i)})^2}{\delta_i^2}, \qquad b_k(\theta) = \sum_{\{i : f_i(x,\theta)<1\}} \frac{\beta_k^{(i)}}{\delta_i^2}$$

$$c_{kl}(\theta) = \sum_{\{i : f_i(x,\theta)<1\}} \frac{\beta_k^{(i)} \beta_l^{(i)}}{\delta_i^2}, \qquad f_i(x,\theta) = \sum_{k=1}^{M} \theta_k \beta_k^{(i)}$$

Main Computational Ingredient

Define the following linear system:

$$
\underbrace{\begin{pmatrix} a_1 & \dots & c_{1M} \\ \vdots & \ddots & \vdots \\ c_{1M} & \dots & a_M \end{pmatrix}}_{C \in \mathbb{R}^{M \times M}} \underbrace{\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_M \end{pmatrix}}_{\theta \in \mathbb{R}^M} = \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_M \end{pmatrix}}_{b \in \mathbb{R}^M} \tag{8}
$$

where $C = C(\theta)$ and $b = b(\theta)$. This system defines a fixed point mapping:

$$
\Psi : \mathbb{R}^M \to \mathbb{R}^M
$$

with

$$
\theta = \Psi(\theta) = C(\theta)^{-1} b(\theta) \tag{9}
$$

Theorem
Let $B = [\beta_1, \dots, \beta_M]$ *be a factor matrix, such that each* $\beta_j \in \mathbb{R}^p$, $p > k$ *and* $\Delta = diag(\delta_1^2, \dots, \delta_p^2)$. *For a structured matrix of the form* $\Sigma = BB^\top + \Delta$, *the solution to the Long-Only Markowitz problem is given by:*

$$x_i = \frac{\left(1 - \sum_{k=1}^{M} \theta_k \beta_k^{(i)}\right)_+}{Z\delta_i^2}, \quad Z = \sum_{i=1}^{p} \frac{\left(1 - \sum_{k=1}^{M} \theta_k \beta_k^{(i)}\right)_+}{\delta_i^2} \tag{10}$$

which can be computed as the solution to the fixed-point mapping

$$\theta = \Psi(\theta) = C(\theta)^{-1} b(\theta) \qquad \theta \in \mathbb{R}^M \tag{11}$$

Sketch of Proof

Idea of Proof:

- Use KKT Conditions on Lagrangian with implicit inclusion of positivity constraint by squaring; i.e. $x_i = v_i^2$; define v^2 as the vector of v_i^2 elements.
- We can rewrite the long-only minimization problem as:

$$\min_{\|v\|_2=1} (v^2)^\top \Sigma v^2$$

- θ_k represents a scaling vector for β_k; If the elements of β_k get too big, they get set to 0 by our positivity requirement.

Fixed Point Algorithm

We solve the following fixed point:

$$\theta = \Psi(\theta)$$
$$\theta = C(\theta)^{-1} b(\theta)$$

This can be computed as follows:

1. Initialize θ^0 as some initial condition; Define ε as some small tolerance.
2. Iterate $\theta^{n+1} = C(\theta^n)^{-1} b(\theta^n)$
   (or equivalently, solve $C(\theta^n)\theta^{n+1} = b(\theta^n)$)
3. Terminate when $\left|\theta^{n+1} - \theta^n\right| < \varepsilon$.

Advantages/Comments

- Interpretability of the Solution:
  - Portfolio Composition
  - Sensitivity to Parameters
  - (semi)-explicit formulae are useful, in general
- Note that factor variances ($\sigma_k^2, k \in \{1 \ldots, M\}$) have very little effect; only appear in the term a_k and are relatively unimportant.
- Efficiency (can compute optimal very quickly)- solving for θ is now a function of M (number of factors) and not p (number of equities)- can lead to very fast recalculation of portfolios.
- Actual convergence rate and convergence guarantees are still an open question we are working on.

Computational Speed

Portfolio Size	Explicit Average Time (ms)	Numeric Average Time (ms)	Time Ratio (Numeric/Explicit)
256	0.341	17.942	5.27×10^1
512	0.444	103.060	2.32×10^2
1024	1.644	1128.559	6.86×10^2
2048	0.932	2757.903	2.96×10^3
4096	1.521	21930.225	1.44×10^4

- All portfolio optimizations repeated and averaged across 100 trials. Randomized 6-factor model was used.
- (L^2) norm differences in portfolio values at termination on the order of 10^{-9} for each example.

Summary and Conclusions:

- Now have semi-explicit formulas for solutions to long-only Markowitz Portfolios under factor-model constraints
  - Mathematically analyze portfolio sensitivity to parameters
  - Can study geometry of the solution
- Significant efficiency gains for portfolio calculations of these models

Future Work

Future Areas to Explore:

- Generalize to coordinate constrained programming for quadratic optimization
- Use of Random Matrix Theory to understand distributions of portfolio weights
- Efficiency gains potentially allow for recalculation of portfolios in milliseconds, which could make it usable in High-Frequency Trading

Clarke, Roger, Harindra de Silva & Steven Thorley (2011), 'Minimum-variance portfolio composition', *The Journal of Portfolio Management* **37**(2), 31–45.
URL: *https://jpm.iijournals.com/content/37/2/31*