Analytic Solutions to the Markowitz Problem.

December 1, 2021.

1. Introduction

In the Capital Asset Pricing Model, every market investor solves a quadratic optimization of the form

$$\begin{aligned} &\min_{x \in \mathbf{R}^p} \langle x, \boldsymbol{\Sigma} x \rangle \\ &\langle x, \alpha \rangle \geq \mu \\ &\langle x, \mathrm{e} \rangle = 1. \end{aligned} \tag{1}$$

An explicit solution x^* can be easily obtained under this formulation. However, Markowitz originally envisioned this problem with the additional constraint

$$x \geq 0.$$

No analytical formula is known.

2. A simplified formulation

To simplify this problem, we make the assumption that

$$\Sigma = \sigma^2 \beta \beta^\top + \Delta, \tag{2}$$

and consider problem of the form

$$\begin{aligned} &\min_{x \in \mathbf{R}^p} \langle x, \boldsymbol{\Sigma} x \rangle \\ &\langle x, v \rangle = 1 \\ &x \geq 0. \end{aligned} \tag{3}$$

Theorem 2.1. *The (unique) solution of (3) is given by $x = \frac{w}{\langle w, \mathrm{v} \rangle}$ where $w = \Delta^{-1}(\mathrm{v} - \beta\theta)_+$ for θ the (unique) fixed point of*

$$\psi(\vartheta) = \frac{\sum_{\vartheta \beta_i < v_i} \beta_i / \delta_i^2}{1/\sigma^2 + \sum_{\vartheta \beta_i < v_i} \beta_i^2 / \delta_i^2} \tag{4}$$

3. Finding the fixed point

Surprisingly, a simple fixed point iteration of this map ψ starting at $\theta_0 = 0$ can give us the fixed point in p steps. This relies on some special properties of the map. The rest of the section consists of its proof.

We have the following built in assumptions from the problem setup:

Assumptions.

1. $\psi : \mathbb{R} \to \mathbb{R}$
2. $\psi(0) > 0$.
3. $v \in \mathbb{R}^p_+$.

For convenience of writing, we use the following notations:

Notations.

1. $\frac{v_{-m}}{\beta_{-m}} < ... < \frac{v_{-1}}{\beta_{-1}} < 0 < \frac{v_1}{\beta_1} < ... < \frac{v_n}{\beta_n}$, with $n + m = p$.
2. $x_k := \sum_{i=k}^n v_i \beta_i$, $y_k := \frac{1}{\sigma^2} + \sum_{i=k}^n \beta_i^2$, for $1 \le k \le n$.
3. x^* is the fixed point (which we prove will always exist and is unique)

From now on, all sub-indices will be positive since we are restricting our attention to the value of psi on the positive axis only.

Lemma. There exists a unique interval $I_i := [\frac{v_i}{\beta_i}, \frac{v_{i+1}}{\beta_{i+1}})$ s.t. $x^* \in I$ and ψ is attains its maximum on I.

Proof. Pick arbitrary $\sigma \neq 0$. Without loss of generality assume $\delta = (1, ..., 1)$, and β_i distinct (merge otherwise).

Claim 1. $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}})$ iff $\psi(\frac{v_i}{\beta_i}) < \frac{v_{i+1}}{\beta_{i+1}}$.

> Condition that characterizes whether psi will increase or decrease, going from $\frac{v_i}{\beta_i}$ to $\frac{v_{i+1}}{\beta_{i+1}}$.

Proof of claim 1. We have

$$\frac{x_i - v_{i+1}\beta_{i+1}}{y_i - \beta_{i+1}^2} < \frac{x_i}{y_i} \iff \frac{x_i}{y_i} < \frac{v_{i+1}}{\beta_{i+1}}$$

Note that we have utilized the fact $y_i > 0, y_i - \beta_{i+1}^2 > 0$.

Claim 2. $\psi(\frac{v_i}{\beta_i}) < \psi(\frac{v_{i+1}}{\beta_{i+1}})$ iff $\psi(\frac{v_i}{\beta_i}) > \frac{v_i}{\beta_i}$.

Rule for telling whether $\psi(x) > x$ or $\psi(x) < x$.

Proof of claim 2.
By claim 1., $\psi(\frac{v_i}{\beta_i}) < \psi(\frac{v_{i+1}}{\beta_{i+1}})$ implies $\psi(\frac{v_i}{\beta_i}) > \frac{v_{i+1}}{\beta_{i+1}}$. We then have

$$\psi(\frac{v_i}{\beta_i}) > \frac{v_{i+1}}{\beta_{i+1}} > \frac{v_i}{\beta_i}.$$

The implication in the reverse direction follows in the same way.

Claim 3. If $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}})$, then $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}}) > ... > \psi(\frac{v_n}{\beta_n})$.

Decrease for once means decrease forever since.

Proof of claim 3. Assume $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}})$, then we get

$$\psi(\frac{v_{i+1}}{\beta_{i+1}}) < \psi(\frac{v_i}{\beta_i}) < \frac{v_i}{\beta_i} < \frac{v_{i+1}}{\beta_{i+1}},$$

where the second inequality follows from the fact that $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}})$ implies $\psi(\frac{v_i}{\beta_i}) < \frac{v_i}{\beta_i}$ proven in claim 2.
Together, claim 2. and equation (3) implies $\psi(\frac{v_{i+2}}{\beta_{i+2}}) < \psi(\frac{v_{i+1}}{\beta_{i+1}})$. The claim then follows from induction.

Claim 4. If $x^* \in [\frac{v_i}{\beta_i}, \frac{v_{i+1}}{\beta_{i+1}})$ and $1 < i < p$, then $\psi(\frac{v_{j-1}}{\beta_{j-1}}) < \psi(\frac{v_j}{\beta_j})$ and $\psi(\frac{v_j}{\beta_j}) > \psi(\frac{v_{j+1}}{\beta_{j+1}})$.
Proof. Direct consequence of claim 2.

Now we go back to prove the **lemma**.
Case 1) ψ is monotonically non-decreasing on $[0, \infty)$. By claim 4, $\psi(\frac{v_n}{\beta_n}) > \frac{v_n}{\beta_n}$, and $\psi(x) = \psi(\frac{v_n}{\beta_n})$ for all $x \in [\frac{v_n}{\beta_n}, \infty)$. So ψ must cross the identity function exactly once at its maximal value, which gives us the desired claim.
Case 2) ψ is not monotonically non-decreasing, i.e., there exists a j where $\psi(\frac{v_i}{\beta_i}) > \psi(\frac{v_{i+1}}{\beta_{i+1}})$. Let j be the first instance where this occurs, then $\psi(\frac{v_{j-1}}{\beta_{j-1}}) < \psi(\frac{v_j}{\beta_j})$ and $\psi(\frac{v_j}{\beta_j}) > \psi(\frac{v_{j+1}}{\beta_{j+1}})$ by claim 4.. By claim 2., we have $\psi(\frac{v_j}{\beta_j}) > \frac{v_j}{\beta_j}$ and $\psi(\frac{v_{j+1}}{\beta_{j+1}}) < \frac{v_{j+1}}{\beta_{j+1}}$. Thus, $x^* \in [\frac{v_j}{\beta_j}, \frac{v_{j+1}}{\beta_{j+1}})$. By claim 3., $\psi(\frac{v_j}{\beta_j}) > \psi(\frac{v_{j+1}}{\beta_{j+1}}) > ... > \psi(\frac{v_n}{\beta_n})$, so x^* is the unique maximizer.

This completes the proof of the lemma.

Corollary. $\psi^{(n)}(0)$ *converges to* x^* *in at most* p *steps.*

Proof. If $\psi(0) = 0$, we are done in one step. So without loss of generality we assume $\psi(0) > 0$. Then there are two possible scenarios:
a) $\psi(x_0)$ is the fixed point. Done.
b) $\psi(x_0)$ is not the fixed point. Then $\psi(x_0) > \frac{1}{\beta_1}$ and $\psi^{(2)}(x_0) > \psi(x_0)$. So $\psi(x_0)$ belongs to an interval $[\frac{1}{\beta_j}, \frac{1}{\beta_{j+1}})$ that is different from the one which $\psi(x_0)$ belongs to.
Inductively, we know that if $\psi^{(n-1)}(x_0) \neq x^*$, then

$$\psi^{(n-1)}(x_0) > \psi^{(n-2)}(x_0) > ... > \psi^{(0)}(x_0) = x_0$$

Thus $\psi^{(p)}(x_0) = x^*$.

4. Adding one inequality constraint

We extend the formula of Clarke et. al. (2011) to a mean-variance optimization setting. In particular, for a single-index model covariance matrix $\mathbf{\Sigma} = \sigma^2 \beta\beta^\top + \mathbf{\Delta}$ we provide an *analytical* formula for

$$
\begin{aligned}
&\min_{x \in \mathbf{R}^p} \langle x, \mathbf{\Sigma} x \rangle \\
&\langle x, \alpha \rangle \geq \mu \\
&\langle x, \mathrm{e} \rangle = 1 \\
&x \geq 0,
\end{aligned}
\tag{5}
$$

where $\mathrm{e} = (1, \ldots, 1)$ and α are both in $\mathbf{R}^p$. We assume α and e are linearly independent and that $\mu \leq \max_i |\alpha_i|$ (otherwise no solution). Let $\langle u, v \rangle = \sum_i u_i v_i$. The solution x is given by $\frac{w}{\langle w, \mathrm{e} \rangle}$ for

$$
w = \mathbf{\Delta}^{-1}(\alpha + \rho \mathrm{e} - \theta\beta)_+ \tag{6}
$$

where (θ, ρ) is the unique fixed point $= \psi()$ for ψ satisfying

$$
\begin{aligned}
\psi(\omega) &= \mathbf{A}_\omega^{-1} \mathrm{b}_\omega, \\
\langle u, v \rangle_{\mathbf{W}} &= \sum_{\omega_1 \beta_i - \omega_2 < \alpha_i} u_i v_i / \delta_i^2, \\
\mathbf{A}_\omega &= \begin{pmatrix} 1/\sigma^2 + \langle \beta, \beta \rangle_{\mathbf{W}} & -\langle \mathrm{e}, \beta \rangle_{\mathbf{W}} \\ \langle \beta, \alpha \rangle_{\mathbf{W}} - \mu \langle \beta, \mathrm{e} \rangle_{\mathbf{W}} & \mu \langle \mathrm{e}, \mathrm{e} \rangle_{\mathbf{W}} - \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}} \end{pmatrix}, \\
\mathrm{b}_\omega &= \begin{pmatrix} \langle \alpha, \beta \rangle_{\mathbf{W}} \\ \langle \alpha, \alpha \rangle_{\mathbf{W}} - \mu \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}} \end{pmatrix}.
\end{aligned}
$$

5. Proof

Consider the change of variables $y_i^2 = x_i$ and define,

$$G(y) = \sum_{i=1}^{p} y_i^2 \beta_i = \sum_{y_i^2 \neq 0} y_i^2 \beta_i . \tag{7}$$

In this direction, we define a weighed norm $\langle \cdot, \cdot \rangle_{\mathbf{W}}$ by

$$\langle u, v \rangle_{\mathbf{W}} = \sum_{y_i^2 \neq 0} u_i v_i / \delta_i^2 . \tag{8}$$

The Lagrangian for the problem is given by

$$\mathscr{L}(v, \ell, \eta) = \sigma^2 G^2(v) + \langle \delta^2, v^4 \rangle + 2\ell\big(\mu - \langle \alpha, v^2 \rangle\big) + 2\eta\big(1 - \langle v^2, \mathrm{e} \rangle\big) .$$

The KKT conditions read,

$$\begin{aligned} 4\sigma^2 G(v)\beta_i v_i + 4\delta_i^2 v_i^3 - 4\ell v_i \alpha_i - 4\eta v_i &= 0 \\ \langle v^2, \alpha \rangle &= \mu \text{ or } \ell = 0 \\ \langle v^2, \mathrm{e} \rangle &= 1 \end{aligned}$$

We have, either $x_i = v_i^2 = 0$ or

$$x_i = v_i^2 = \frac{\eta + \ell \alpha_i - \sigma^2 G(v) \beta_i}{\delta_i^2} \tag{9}$$

$$= \Big(\frac{\ell}{\delta_i^2}\Big)\big(\rho + \alpha_i - \theta \beta_i\big) > 0 \tag{10}$$

where $\rho = \eta/\ell$ ($\ell > 0$ if the μ-constraint is binding) and $\theta = \sigma^2 G(v)/\ell$. Define $\langle a, b \rangle_{\mathbf{W}} = \sum_{v_i^2 \neq 0} a_i b_i / \delta_i^2$. We obtain the following equations.

$$\begin{aligned} \theta &= \sigma^2\big(\rho \, \langle \mathrm{e}, \beta \rangle_{\mathbf{W}} + \langle \alpha, \beta \rangle_{\mathbf{W}} - \theta \langle \beta, \beta \rangle_{\mathbf{W}}\big) \\ \mu &= \ell\big(\rho \, \langle \mathrm{e}, \alpha \rangle_{\mathbf{W}} + \langle \alpha, \alpha \rangle_{\mathbf{W}} - \theta \langle \beta, \alpha \rangle_{\mathbf{W}}\big) \text{ (or } \ell = 0) \\ 1 &= \ell\big(\rho \, \langle \mathrm{e}, \mathrm{e} \rangle_{\mathbf{W}} + \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}} - \theta \langle \beta, \mathrm{e} \rangle_{\mathbf{W}}\big) \end{aligned}$$

When $\ell > 0$ and $\big(\rho \, \langle \mathrm{e}, \mathrm{e} \rangle_{\mathbf{W}} + \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}} - \theta \langle \beta, \mathrm{e} \rangle_{\mathbf{W}}\big) > 0$ we can eliminate it to reduce to a system in (ρ, θ).

$$\begin{aligned} \big(1/\sigma^2 + \langle \beta, \beta \rangle_{\mathbf{W}}\big)\, \theta - \langle \mathrm{e}, \beta \rangle_{\mathbf{W}} \rho &= \langle \alpha, \beta \rangle_{\mathbf{W}} \\ \big(\langle \beta, \alpha \rangle_{\mathbf{W}} - \mu \langle \beta, \mathrm{e} \rangle_{\mathbf{W}}\big)\, \theta + \big(\mu \langle \mathrm{e}, \mathrm{e} \rangle_{\mathbf{W}} - \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}}\big)\, \rho &= \langle \alpha, \alpha \rangle_{\mathbf{W}} - \mu \langle \alpha, \mathrm{e} \rangle_{\mathbf{W}} \end{aligned}$$

We leave the discussion of the case when $\big(\rho\langle \mathrm{e},\mathrm{e}\rangle_{\mathbf{W}} + \langle \alpha,\mathrm{e}\rangle_{\mathbf{W}} - \theta\langle \beta,\mathrm{e}\rangle_{\mathbf{W}}\big) = 0$ to the next section. Consequently, we have the linear system $\mathbf{A} = \mathrm{b}$ for $= (\theta,\rho)$ and

$$\mathbf{A} = \begin{pmatrix} 1/\sigma^2 + \langle \beta,\beta\rangle_{\mathbf{W}} & -\langle \mathrm{e},\beta\rangle_{\mathbf{W}} \\ \langle \beta,\alpha\rangle_{\mathbf{W}} - \mu\langle \beta,\mathrm{e}\rangle_{\mathbf{W}} & \mu\langle \mathrm{e},\mathrm{e}\rangle_{\mathbf{W}} - \langle \alpha,\mathrm{e}\rangle_{\mathbf{W}} \end{pmatrix},$$
$$\mathrm{b} = \begin{pmatrix} \langle \alpha,\beta\rangle_{\mathbf{W}} \\ \langle \alpha,\alpha\rangle_{\mathbf{W}} - \mu\langle \alpha,\mathrm{e}\rangle_{\mathbf{W}} \end{pmatrix}.$$

Both $\mathbf{A}$ and b depend on $= (\theta,\rho)$ since, for example,

$$\langle \alpha,\beta\rangle_{\mathbf{W}} = \sum_{v_i^2\neq 0} \alpha_i\beta_i/\delta_i^2 = \sum_{\theta\beta_i-\rho<\alpha_i} \alpha_i\beta_i/\delta_i^2. \tag{11}$$

This leads to the definition of a fixed point equation $= \psi()$ where

$$\psi(\omega) = \mathbf{A}_\omega^{-1}\mathrm{b}_\omega$$

Once the fixed point $= (\theta,\rho)$ is determined, we have

$$x = \frac{w}{\langle w,\mathrm{e}\rangle} \tag{12}$$

where $w = \mathbf{\Delta}^{-1}(\alpha + \rho\mathrm{e} - \theta\beta)_+$.

6. Bad fixed points

As mentioned in the previous section, when turning the 3 by 3 system of equations to the 2 by 2 system, the condition $\big(\rho\langle \mathrm{e},\mathrm{e}\rangle_{\mathbf{W}} + \langle \alpha,\mathrm{e}\rangle_{\mathbf{W}} - \theta\langle \beta,\mathrm{e}\rangle_{\mathbf{W}}\big) > 0$ is not always satisfied. This happens if and only if

$$\rho \leq \min_{1\leq i\leq p} (\theta\beta_i - \alpha_i).$$

Applying this restriction on ρ to the equation

$$\theta = \sigma^2\big(\rho\langle \mathrm{e},\beta\rangle_{\mathbf{W}} + \langle \alpha,\beta\rangle_{\mathbf{W}} - \theta\langle \beta,\beta\rangle_{\mathbf{W}}\big),$$

we get $\theta = 0$, meaning $\{(0,\rho) : \rho \leq \min_{1\leq i\leq p} -\alpha_i\}$ contains the possible set of solutions to the 2 by 2 system that does not solve the 3 by 3 system. One can easily verify that this is exactly the set of bad fixed points.

7. Brute force algorithm

The function $\psi(w) = A_w^{-1} b_w$ has the following property: it remains constant inside each of the disjoint region on the (θ, ρ) plane partitioned by the equations $\rho = \beta_i \theta - \alpha_i$ for $1 \leq i \leq n$. Let x^* denote the correct fixed point and suppose it lies in region D. Then by the property of ψ, we know that $y \in D$ if and only if $\psi(y) = x^*$ and $\psi(\psi(y)) = x^*$. Naturally, one can identify the fixed points by brute forcing all possible disjoint regions, whose total number belongs to $O(p^2)$.

Algorithm 1 Brute force

Precondition: $\alpha, \beta, \delta, \sigma$ are parameters from the Markovitz model; L is a list that store the candidate solutions of the form (θ, ρ); Σ is the covariance matrix defined as $\Sigma := \sigma^2 \beta \beta^\top + \Delta$.

function BruteForce($\alpha, \beta, \delta, \sigma$)
  for $i \leftarrow 1$ to n **do**
    for $j \leftarrow 1$ to n **do**
      $\theta_{ij} \leftarrow$ solution to $\beta_i \theta - \alpha_i = \beta_j \theta - \alpha_j$ ▷ $\theta_{ij} \in \mathbb{R}^2$
      $g_i \leftarrow (-\beta_i, 1)$
      $g_j \leftarrow (-\beta_j, 1)$
      $(s_1, s_2, s_3, s_4) \leftarrow (g_i + g_j, g_i - g_j, -g_i + g_j, -g_i - g_j)$
      for $k \leftarrow 1$ to 4 **do**
        $y_1 \leftarrow \psi(\theta_{ij} + s_i)$
        $y_2 \leftarrow \psi(y_1)$
        if $y_1 = y_2$ and $\theta \neq 0$ **then**
          return $\alpha + \rho e - \theta \beta$
        end if
      end for
    end for
  end for
end function

Claim. The worst-case runtime is $O(p^2)$.
Proof. We first compute $r(n)$, the maximum total number of disjoint regions that can be partitioned by n straight lines in the two dimensional plane. Inductively, one can check that $r(n) = r(n-1) + n$, and thus $r(p) \in O(p^2)$.
In the worst case, the brute force algorithm arrives at the correct region after checking all the incorrect regions. At each region there are two newton updates with constant cost. Therefore, the worst case complexity of the brute force algorithm is simply $O(p^2)$.

In practice, it is helpful to randomizes the order (uniformly) of the alpha and beta parameters to prevent an adversarial ordering of those parameters. To analyze its performance, we define N to be the total number of regions and X the number of trials before arriving at the correct region. Since $\mathrm{P}(X = d) = \frac{N-1}{N}\frac{N-2}{N-1}\cdots\frac{1}{N-d} = \frac{1}{N}$, we get $\mathrm{E}(X) \in O(N)$.
Recall that $N(p) \in O(p^2)$. So

$$\mathrm{E}(X) \in O(p^2).$$

Thus, the average complexity remains on the same order as the worst-case complexity.