Math 5760/6890: Introduction to Mathematical Finance

The N-security Markowitz Portfolio

See Petters and Dong 2016, Section 3.3-3.4

Akil Narayan[1]

[1]Department of Mathematics, and Scientific Computing and Imaging (SCI) Institute
University of Utah

Fall 2024

A quick recap

For Markowitz 2-security portfolio optimization:

- Return rates $\boldsymbol{R}$ for the two securities are random variables
- Assume first- and second-order statistics of these are available: $\boldsymbol{\mu} = \mathbb{E}\boldsymbol{R}$ and $\boldsymbol{A} = \mathrm{Cov}(\boldsymbol{R})$
- The portfolio is defined by the weights $\boldsymbol{w}$
- The expected return rate of the portfolio is $\mu_P = \langle \boldsymbol{\mu}, \boldsymbol{w} \rangle$
- The squared risk of the portfolio is $\sigma_P^2 = \mathrm{Var}\langle \boldsymbol{R}, \boldsymbol{w} \rangle = \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w}$

A quick recap

For Markowitz 2-security portfolio optimization:

- Return rates $\boldsymbol{R}$ for the two securities are random variables
- Assume first- and second-order statistics of these are available: $\boldsymbol{\mu} = \mathbb{E}\boldsymbol{R}$ and $\boldsymbol{A} = \mathrm{Cov}(\boldsymbol{R})$
- The portfolio is defined by the weights $\boldsymbol{w}$
- The expected return rate of the portfolio is $\mu_P = \langle \boldsymbol{\mu}, \boldsymbol{w} \rangle$
- The squared risk of the portfolio is $\sigma_P^2 = \mathrm{Var}\langle \boldsymbol{R}, \boldsymbol{w} \rangle = \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w}$
- Generally, no optimization is needed for 2-security portfolios: the weight constraints uniquely identify portfolios
- Many risk-optimal portfolios are *not* efficient
- The set of risk-optimal portfolios can be explicitly parameterized + plotted
- There is a global variance-minimizing portfolio – it's possible investors might not want this portfolio.
- The *efficient frontier* can be "easily" identified.

We now consider the risk-minimization problem for an N-security portfolio:

$$\begin{aligned} \min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w} \;\; \text{subject to} \;\; & \langle \boldsymbol{w}, \mathbf{1} \rangle = 1, \;\text{and} \\ & \langle \boldsymbol{w}, \boldsymbol{\mu} \rangle = \mu_P. \end{aligned}$$

We now consider the risk-minimization problem for an N-security portfolio:

$$\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w} \;\; \text{subject to} \;\; \langle \boldsymbol{w}, \boldsymbol{1} \rangle = 1, \;\text{and}$$
$$\langle \boldsymbol{w}, \boldsymbol{\mu} \rangle = \mu_P.$$

- The constraints are the same: we enforce $\boldsymbol{w}$ is a valid portfolio weight, and that the portfolio has expected return rate μ_P.
- Just like before, we will see that risk-optimal portfolios need not be efficient ones.
- We make assumptions that are, by now, commonplace:
  - $\boldsymbol{A}$ is positive-definite
  - $\boldsymbol{\mu}$ is not parallel to $\boldsymbol{1}$
  - We allow short selling

Computing the risk-optimal portfolio

$$\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w} \quad \text{subject to} \quad \langle \boldsymbol{w}, \boldsymbol{1} \rangle = 1, \text{ and}$$
$$\langle \boldsymbol{w}, \boldsymbol{\mu} \rangle = \mu_P.$$

We assume $\boldsymbol{A}$ is positive-definite and $\boldsymbol{\mu}$ is not parallel to $\boldsymbol{1}$.

Unlike the 2-security case, we have a real optimization problem: the two linear constraints are not sufficient to uniquely characterize a vector with 3 or more entries.

To solve this constrained optimization problem, we'll use (+recall) Lagrange multipliers.

Recall: when optimizing a smooth function $f(\underline{x})$, $\underline{x} \in \mathbb{R}^N$, without constraints: $\nabla f(\underline{x}) = \underline{0}$ is a necessary condition for optimality.

If we have equality constraints, say $g(\underline{x}) = 0$ (single constraint), then a necessary condition for optimality (subject to constraint) is

$$\nabla f(\underline{x}) = \lambda \nabla g(\underline{x})$$
$$g(\underline{x}) = 0$$

(for some scalar λ)

More general case of Lagrange multipliers: optimize $f(\underline{x})$ subject to M constraints: $g_1(\underline{x}) = 0$, $g_2(\underline{x}) = 0, \ldots$ $g_M(\underline{x}) = 0$

Compute stationary points: $L(\underline{x}, \underline{\lambda}) = f(\underline{x}) - \sum_{j=1}^{M} \lambda_j (g_j(\underline{x}) - 0)$

L: "Lagrangian" $\qquad \underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_M \end{pmatrix} \in \mathbb{R}^n$

$$L: \mathbb{R}^{N+M} \to \mathbb{R}$$
$$f: \mathbb{R}^N \to \mathbb{R}$$
$$g_j: \mathbb{R} \to \mathbb{R} \text{ for all } j = 1, \ldots M$$

Stationary points: $\nabla L(\underline{x}, \underline{\lambda}) = \underline{0}$ (This means simultaneously $\nabla_{\underline{x}} L = 0$, $\nabla_{\underline{\lambda}} L = 0$)

$$\Rightarrow \nabla f - \sum_{j=1}^{M} \lambda_j \nabla g_j(\underline{x}) = \underline{0} \in \mathbb{R}^N$$
$$g_1(\underline{x}) = 0$$
$$g_2(\underline{x}) = 0 \; \ldots \; g_M(\underline{x}) = 0$$

(we'll assume stationary points are local extrema)

Back to our problem: $\min \underline{w}^T \underline{\underline{A}}\, \underline{w}$ subject to $\underline{w}^T \underline{1} = 1$

$\underline{w}^T \underline{\mu} = \mu_p$

$\underline{\underline{A}}, \underline{\mu}, \mu_p$ given.

$f(\underline{w}) = \underline{w}^T \underline{\underline{A}}\, \underline{w}, \quad g_1(\underline{w}) = \underline{w}^T \underline{1} - 1, \quad g_2(\underline{w}) = \underline{w}^T \underline{\mu} - \mu_p.$

$\min_{\underline{w}} f(\underline{w})$ s.t. $g_1(\underline{w}) = 0$, $g_2(\underline{w}) = 0$ "M" = 2

$\underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$

Lagrangian: $L(\underline{w}, \underline{\lambda}) = \underline{w}^T \underline{\underline{A}}\, \underline{w} - \lambda_1 (\underline{w}^T \underline{1} - 1) - \lambda_2 (\underline{w}^T \underline{\mu} - \mu_p)$

$\nabla_{\underline{w}} L = \underline{0} \implies 2\underline{\underline{A}}\, \underline{w} - \lambda_1 \underline{1} - \lambda_2 \underline{\mu} = \underline{0}$ (1)

$\nabla_{\underline{\lambda}} L = 0 \implies \underline{1}^T \underline{w} = 1$ (2)

$\underline{\mu}^T \underline{w} = \mu_p$ (3)

(1): solve for $\underline{w}$: $\underline{w} = \frac{1}{2} [\lambda_1 \underline{\underline{A}}^{-1} \underline{1} + \lambda_2 \underline{\underline{A}}^{-1} \underline{\mu}]$

(2): plug in for $\underline{w}$: $\lambda_1 (\underline{1}^T \underline{\underline{A}}^{-1} \underline{1}) + \lambda_2 (\underline{1}^T \underline{\underline{A}}^{-1} \underline{\mu}) = 2$

$\lambda_1 (\underline{\mu}^T \underline{\underline{A}}^{-1} \underline{1}) + \lambda_2 (\underline{\mu}^T \underline{\underline{A}}^{-1} \underline{\mu}) = 2\mu_p$

$a = \underline{1}^T \underline{\underline{A}}^{-1} \underline{1}$ $\quad a\lambda_1 + b\lambda_2 = 2$

$b = \underline{1}^T \underline{\underline{A}}^{-1} \underline{\mu}$ $\quad b\lambda_1 + c\lambda_2 = 2\mu_p$

$c = \underline{\mu}^T \underline{\underline{A}}^{-1} \underline{\mu}$ $\quad \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ \mu_p \end{pmatrix}$

$$B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad B^{-1} = \frac{1}{ac-b^2}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = B^{-1}\, 2 \begin{pmatrix} 1 \\ \mu_p \end{pmatrix} = \frac{2}{ac-b^2}\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}\begin{pmatrix} 1 \\ \mu_p \end{pmatrix}$$

$$= \frac{2}{ac-b^2}\begin{pmatrix} c - b\mu_p \\ -b + a\mu_p \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$\underline{w} = \frac{1}{2}\left[\lambda_1 \underline{\underline{A}}^{-1} \underline{1} + \lambda_2 \underline{\underline{A}}^{-1} \underline{\mu} \right]$$

$$= \frac{1}{ac-b^2}\left(c\underline{\underline{A}}^{-1}\underline{1} - b\mu_p \underline{\underline{A}}^{-1}\underline{1} - b\underline{\underline{A}}^{-1}\underline{\mu} + a\mu_p \underline{\underline{A}}^{-1}\underline{\mu} \right)$$

(weights corresponding to stationary pt.)

$$= \underline{v}_0 + \mu_p \underline{v}_1, \quad \underline{v}_0 = \frac{c\underline{\underline{A}}^{-1}\underline{1} - b\underline{\underline{A}}^{-1}\underline{\mu}}{ac-b^2}$$

$$\underline{v}_1 = \frac{a\underline{\underline{A}}^{-1}\underline{\mu} - b\underline{\underline{A}}^{-1}\underline{1}}{ac-b^2}$$

$\underline{w} = \underline{v}_0 + \mu_p \underline{v}_1 \Rightarrow \sigma_p^2 = \underline{w}^T \underline{\underline{A}}\, \underline{w}$ is a quadratic function of $\mu_p \Rightarrow$ plot of μ_p vs. σ_p is a hyperbola.

An explicit result

Define some auxilliary quantities:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \mathbf{1}^T \boldsymbol{A}^{-1} \mathbf{1} \\ \mathbf{1}^T \boldsymbol{A}^{-1} \boldsymbol{\mu} \\ \boldsymbol{\mu}^T \boldsymbol{A}^{-1} \boldsymbol{\mu} \end{pmatrix},$$

$$\boldsymbol{v}_0 = \frac{c\boldsymbol{A}^{-1}\mathbf{1} - b\boldsymbol{A}^{-1}\boldsymbol{\mu}}{ac - b^2} \in \mathbb{R}^N, \qquad \boldsymbol{v}_1 = \frac{-b\boldsymbol{A}^{-1}\mathbf{1} + a\boldsymbol{A}^{-1}\boldsymbol{\mu}}{ac - b^2} \in \mathbb{R}^N.$$

An explicit result

Define some auxilliary quantities:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \mathbf{1}^T \boldsymbol{A}^{-1} \mathbf{1} \\ \mathbf{1}^T \boldsymbol{A}^{-1} \boldsymbol{\mu} \\ \boldsymbol{\mu}^T \boldsymbol{A}^{-1} \boldsymbol{\mu} \end{pmatrix},$$

$$\boldsymbol{v}_0 = \frac{c\boldsymbol{A}^{-1}\mathbf{1} - b\boldsymbol{A}^{-1}\boldsymbol{\mu}}{ac - b^2} \in \mathbb{R}^N, \qquad \boldsymbol{v}_1 = \frac{-b\boldsymbol{A}^{-1}\mathbf{1} + a\boldsymbol{A}^{-1}\boldsymbol{\mu}}{ac - b^2} \in \mathbb{R}^N.$$

Theorem

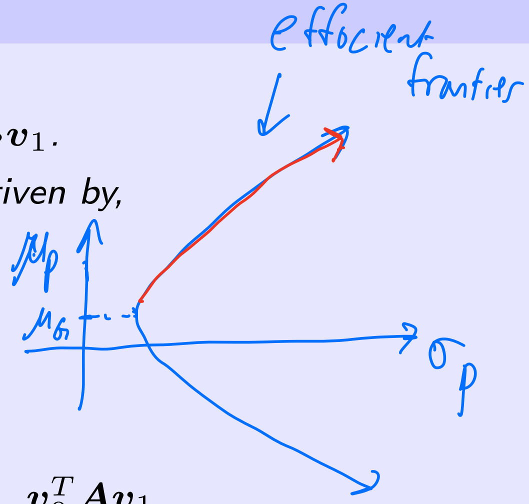
Consider the N-security Markowitz portfolio with our assumptions. Then:

- *The unique risk-optimal portfolio with expected return rate μ_P is $\boldsymbol{w} = \boldsymbol{v}_0 + \mu_P \boldsymbol{v}_1$.*
- *The set of risk-optimal portfolios is a parametric curve in the (σ_P, μ_P) plane, given by,*

$$\frac{\sigma_P^2}{\sigma_G^2} + \frac{(\mu_P - \mu_G)^2}{d^2} = 1.$$

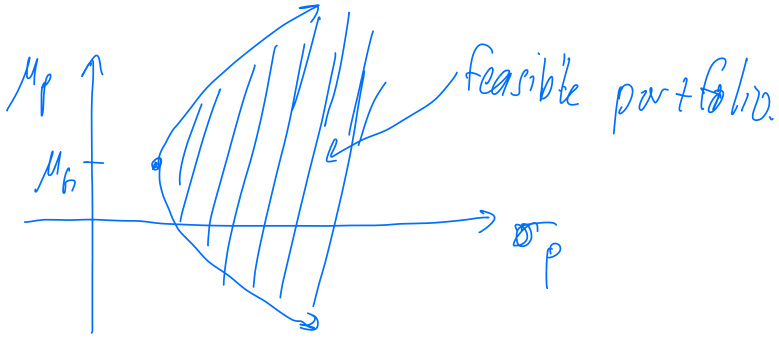
- *The global risk-minimizing portfolio has coordinates (σ_G, μ_G) given by*

$$\sigma_G^2 = \boldsymbol{v}_0^T \boldsymbol{A} \boldsymbol{v}_0 - \frac{(\boldsymbol{v}_0^T \boldsymbol{A} \boldsymbol{v}_1)^2}{\boldsymbol{v}_1^T \boldsymbol{A} \boldsymbol{v}_1}, \qquad \mu_G = -\frac{\boldsymbol{v}_0^T \boldsymbol{A} \boldsymbol{v}_1}{\boldsymbol{v}_1^T \boldsymbol{A} \boldsymbol{v}_1}.$$



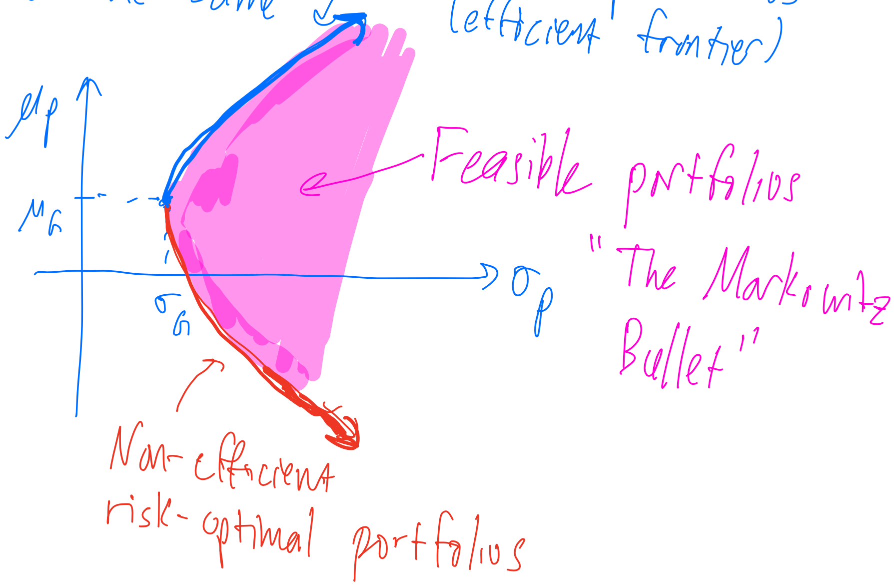
In the 2-security case: the hyperbola is the set of feasible portfolio.

For $N > 2$, this is not true



Portfolios not on the graph of the hyperbola is not risk optimal (is not efficient)

But many things are the same:



The efficient frontier

As with the 2-security case, the efficient frontier corresponds to the portfolios on the "upper half" of the risk-minimizing hyperbola.

I.e., a risk-minimizing portfolio with coordinates (σ_P, μ_P) is *efficient* if $\mu_P \geqslant \mu_G$.

The efficient frontier

As with the 2-security case, the efficient frontier corresponds to the portfolios on the "upper half" of the risk-minimizing hyperbola.

I.e., a risk-minimizing portfolio with coordinates (σ_P, μ_P) is *efficient* if $\mu_P \geqslant \mu_G$.

In the N-security case, there are many more *feasible* portfolios than simply the risk-optimal ones.

In particular, the region of feasible portfolios in the (σ_P, μ_P) plane is the region contained within the hyperbola.

The efficient frontier

As with the 2-security case, the efficient frontier corresponds to the portfolios on the "upper half" of the risk-minimizing hyperbola.

I.e., a risk-minimizing portfolio with coordinates (σ_P, μ_P) is *efficient* if $\mu_P \geqslant \mu_G$.

In the N-security case, there are many more *feasible* portfolios than simply the risk-optimal ones.

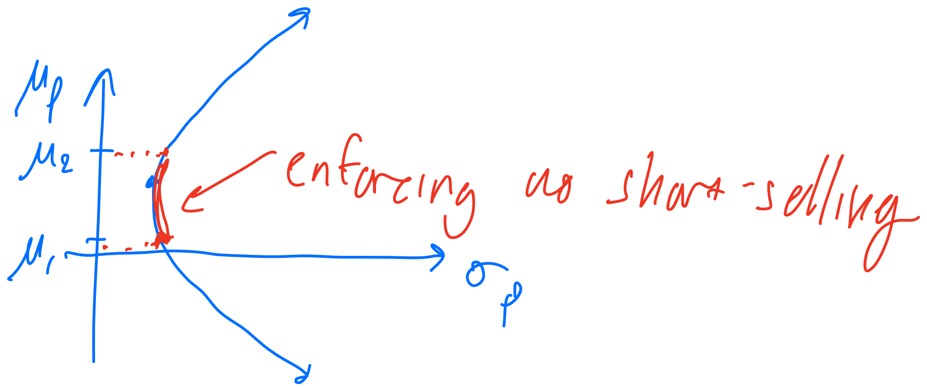
In particular, the region of feasible portfolios in the (σ_P, μ_P) plane is the region contained within the hyperbola.

This region of feasible portfolios is called the *Markowitz bullet*, with the risk-optimal portfolios forming the bullet boundary, the efficient frontier the upper part of this boundary, and the global risk-minimizing portfolio the tip of the bullet.

Efficient portfolios can correspond to substantially leveraged portfolios.

What if we disallow short selling? Things get kind of complicated.

Recall in 2-security case:



Portfolios without short selling

Efficient portfolios can correspond to substantially leveraged portfolios.

What if we disallow short selling? Things get kind of complicated.

$$\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{A} \boldsymbol{w} \text{ subject to } \langle \boldsymbol{w}, \boldsymbol{1} \rangle = 1, \text{ and}$$
$$\langle \boldsymbol{w}, \boldsymbol{\mu} \rangle = \mu_P, \text{and}$$
$$w_i \geqslant 0, \quad i = 1, 2, \ldots, N.$$

This is an inequality-constrained optimization problem.

The method of Lagrange multipliers can no longer handle this problem. (We need instead the *Karush-Kuhn-Tucker* conditions.)

Efficient portfolios can correspond to substantially leveraged portfolios.

What if we disallow short selling? Things get kind of complicated.

$$\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{A}\boldsymbol{w} \text{ subject to } \begin{aligned} \langle \boldsymbol{w}, \mathbf{1} \rangle &= 1, \text{ and} \\ \langle \boldsymbol{w}, \boldsymbol{\mu} \rangle &= \mu_P, \text{and} \\ w_i &\geqslant 0, \quad i = 1, 2, \ldots, N. \end{aligned}$$

This is an inequality-constrained optimization problem.

The method of Lagrange multipliers can no longer handle this problem. (We need instead the *Karush-Kuhn-Tucker* conditions.)

Typically we rely on numerical software to solve this problem, but generically the solution set is some subset of the Markowitz bullet, with the efficient frontier a subset of the bullet boundary.

References I

Petters, Arlie O. and Xiaoying Dong (2016). *An Introduction to Mathematical Finance with Applications: Understanding and Building Financial Intuition*. Springer. ISBN: 978-1-4939-3783-7.