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

September 19, 2023

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 $\boldsymbol{A} = \mathrm{Cov}(\boldsymbol{R})$ (Book: $V = \mathrm{cov}(R)$)
- The portfolio is defined by the weights $\boldsymbol{w}$

Return rate: $R_p = \langle w, R \rangle$

- 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.

The N-security Markowitz portfolio

Recall: $\sigma_p^2 = w^T A w$

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 N-security Markowitz portfolio

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

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

- 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 $\mathbf{1}$
  - We allow short selling

X, Y : scalar RV's

$$\underline{Z} = \begin{pmatrix} X \\ Y \end{pmatrix}$$

$$Cov(X,Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$$

$$Cov(X,X) = Var(X)$$

$$Cov(\underline{Z}) = \begin{pmatrix} Cov(Z_1, Z_1) & Cov(Z_1, Z_2) \\ Cov(Z_2, Z_1) & Cov(Z_2, Z_2) \end{pmatrix}$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \qquad \sigma_X^2 = Var\,X$$

$$\underline{\underline{\Sigma}}^{(Z)} = diag(\sigma_{Z_1}, \sigma_{Z_2} \ldots \sigma_{Z_N})$$

$$\underline{\underline{\Sigma}}^{-1}\, Cov(Z)\, \underline{\underline{\Sigma}}^{-1} \Leftarrow \text{matrix of correlation coefficients.}$$

Computing the risk-optimal portfolio

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

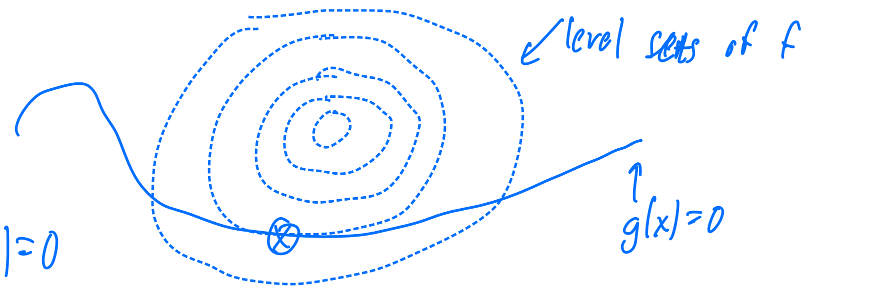
We assume $\boldsymbol{A}$ is positive-definite and $\boldsymbol{\mu}$ is not parallel to $\mathbf{1}$.

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.

Simple version of Lagrange Multipliers: solve the constrained opt. problem: $\min_{\underline{x}} f(\underline{x})$ subject to $g(\underline{x}) = 0$. $\underline{x} \in \mathbb{R}^N$

Plot this problem in $N=2$ dimensions:



at $\otimes$ (solution):

normal vector to $g(x)=0$ and the contour of f are parallel: $\nabla f = \text{const.}\, \nabla g$. (at $\otimes$)

Lagrange multipliers: solving the constraind problem can be achieved by solving:

3 unknowns (x, λ), 3 eqns. $\begin{cases} \nabla f = \lambda \nabla g, & \lambda \in \mathbb{R} \\ g(x) = 0. \end{cases}$

This procedure can be written as an unconstrained opt. problem: Define $\mathcal{L}(\underline{x}, \lambda) = f(x) + \lambda g(x)$

$$\nabla_x \mathcal{L} = 0 \iff \nabla f + \lambda \nabla g = 0$$

$$\nabla_\lambda \mathcal{L} = 0 \iff g(x) = 0$$

$\mathcal{L}$: "Lagrangian".

Lagrange Multiplies: compute $\overbrace{\text{critical}}^{\text{derivative}=0}$ point of $\mathcal{L}$ as a function of $(\underline{x}, \lambda)$

With $m \geq 1$ constraints: $\mathcal{L}(\underline{x}, \underline{\lambda}) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$, $\underline{\lambda} \in \mathbb{R}^m$

where this corresponds to the opt. problem:

$$\min_{x} f(x) \text{ subject to } \underline{g}(\underline{x}) = \underline{0}$$

$$\underline{g}(\underline{x}) = (g_1(\underline{x}), g_2(\underline{x}), \ldots g_m(\underline{x}))^T$$

$$\mathcal{L}(\underline{x}, \underline{\lambda}) = f(\underline{x}) + \underline{g}^T \underline{\lambda}$$

Lagrange multipliers: $\nabla_{\underline{x}, \underline{\lambda}} \mathcal{L} = 0$

$$\Downarrow$$

$$\nabla_{\underline{x}} \rightarrow \nabla f + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0$$

$\underline{\lambda} \in \mathbb{R}^m$

$f: \mathbb{R}^N \to \mathbb{R}$

$$\nabla_{\underline{\lambda}} \begin{cases} g_1(x) = 0 \\ g_2(x) = 0 \\ \vdots \\ g_m(x) = 0 \end{cases}$$

<u>Remark</u>: Critical points (sol'ns to Lagrange mult) need not be minima of f subject to $g(x) = 0$.

Being a solution of a Lagrange mult problem is a necessary condition ~~for~~ for local minimum, but <u>not</u> sufficient.

A sufficient condition can be formulated involving second derivatives of $\mathcal{L}$.

Back to portfolios:

$$\min_{\underline{w}} \underline{w}^T \underline{\underline{A}} \underline{w} \quad \text{subject to} \quad \langle \underline{w}, \underline{1} \rangle = 1$$
$$\langle \underline{w}, \underline{\mu} \rangle = \mu_p$$

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

$$\mathcal{L}(\underline{w}, \underline{\lambda}) = \underbrace{\underline{w}^T \underline{\underline{A}} \underline{w}}_{\sum_{i,j} w_i A_{ij} w_j} + \lambda_1 (\langle \underline{w}, \underline{1} \rangle - 1) + \lambda_2 (\langle \underline{w}, \underline{\mu} \rangle - \mu_p)$$

$$\nabla \mathcal{L} = \underline{0} \implies 2 \underline{\underline{A}} \underline{w} + \lambda_1 \underline{1} + \lambda_2 \underline{\mu} = 0 \quad (1)$$
$$\langle \underline{w}, \underline{1} \rangle = 1 \quad (2)$$
$$\langle \underline{w}, \underline{\mu} \rangle = \mu_p \quad (3)$$

$$(1) \implies \underline{w} = -\tfrac{1}{2} \lambda_1 \underline{\underline{A}}^{-1} \underline{1} - \tfrac{1}{2} \lambda_2 \underline{\underline{A}}^{-1} \underline{\mu}$$

$$(2): \quad \lambda_1 \left[-\frac{1}{2} \underline{1}^T \underline{\underline{A}}^{-1} \underline{1} \right] + \lambda_2 \left[-\frac{1}{2} \underline{1}^T \underline{\underline{A}}^{-1} \underline{\mu} \right] = 1$$

$$(3): \quad \lambda_1 \left[-\frac{1}{2} \underline{\mu}^T \underline{\underline{A}}^{-1} \underline{1} \right] + \lambda_2 \left[-\frac{1}{2} \underline{\mu}^T \underline{\underline{A}}^{-1} \underline{\mu} \right] = \mu_p$$

Define $a = \underline{1}^T \underline{\underline{A}}^{-1} \underline{1}$, $b = \underline{1}^T \underline{\underline{A}}^{-1} \underline{\mu}$, $c = \underline{\mu}^T \underline{\underline{A}}^{-1} \underline{\mu}$

$$(2)+(3): \quad \underbrace{\begin{pmatrix} a & b \\ b & c \end{pmatrix}}_{\underline{\underline{M}}} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -2\mu_p \end{pmatrix}$$

$$\underline{\lambda} = \underline{\underline{M}}^{-1} (-2) \cdot \begin{pmatrix} 1 \\ \mu_p \end{pmatrix}$$

$$= -2 \underline{\underline{M}}^{-1} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_p \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$\underline{\underline{M}}^{-1} = \frac{1}{ac-b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix}$$

$$\Rightarrow \underline{\lambda} = \frac{-2}{ac-b^2} \left[\begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_p \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$$

$$= \frac{-2}{ac-b^2}\left[\begin{pmatrix} c \\ -b \end{pmatrix} + \mu_p \begin{pmatrix} -b \\ a \end{pmatrix}\right]$$

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

Recall: $\underline{w} = -\frac{1}{2}\lambda_1 \underline{\underline{A}}^{-1}\underline{1} - \frac{1}{2}\lambda_2 \underline{\underline{A}}^{-1}\underline{\mu}$

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

$$= \frac{1}{ac-b^2}\left[c\underline{\underline{A}}^{-1}\underline{1} - b\underline{\underline{A}}^{-1}\underline{\mu}\right]$$

$$+ \frac{\mu_p}{ac-b^2}\left[-b\underline{\underline{A}}^{-1}\underline{1} + a\underline{\underline{A}}^{-1}\underline{\mu}\right]$$

$\underline{w}$ ↗ is a critical pt of $\mathcal{I}$ (and we'll assume this corresponds to the global min of the

Opt. problem)

$$\Longrightarrow \underline{v}_0 = \frac{c\,\underline{\underline{A}}^{-1}\underline{1} - b\,\underline{\underline{A}}^{-1}\underline{\mu}}{ac-b^2}$$

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

$$\Longrightarrow \underline{w} = \underline{v}_0 + \mu_p \underline{v}_1$$

Desired (risk-optimal) portfolio.

Q: How does the optimal risk depend on μ_p?

$$\sigma_p^{2\,(\mu_p)} = \underline{w}^T \underline{\underline{A}}\, \underline{w}$$

$$= (\underline{v}_0 + \mu_p \underline{v}_1)^T \underline{\underline{A}} (\underline{v}_0 + \mu_p \underline{v}_1)$$

$$= \mu_p^2 [\underline{v}_1^T \underline{\underline{A}}\, \underline{v}_1] + \mu_p [2 \underline{v}_0^T \underline{\underline{A}}\, \underline{v}_1] + \underline{v}_0^T \underline{\underline{A}}\, \underline{v}_0$$

$$\frac{\sigma_p^2}{v_1^T \underline{\underline{A}} v_1} = \mu_p^2 + \mu_p \frac{2 v_0^T \underline{\underline{A}} v_1}{v_1^T \underline{\underline{A}} v_1} + \frac{v_0^T A v_0}{v_1^T A v_1}$$

$$= \left(\mu_p + \frac{\underline{v}_0^T \underline{\underline{A}} \underline{v}_1}{\underline{v}_1^T \underline{\underline{A}} \underline{v}_1}\right)^2 - \left(\frac{\underline{v}_0^T \underline{\underline{A}} \underline{v}_1}{\underline{v}_1^T \underline{\underline{A}} \underline{v}_1}\right)^2 + \frac{v_0^T A v_0}{v_1^T A v_1}$$

$$\frac{\sigma_p^2}{(v_1^T A v_1)\left[\frac{v_0^T A v_0}{v_1^T A v_1} - \left(\frac{v_0^T A v_1}{v_1^T A v_1}\right)^2\right]} - \frac{\left(\mu_p + \frac{v_0^T A v_1}{v_1^T A v_1}\right)^2}{\frac{v_0^T A v_0}{v_1^T A v_1} - \left(\frac{v_0^T A v_1}{v_1^T A v_1}\right)^2} = 1$$

$$\frac{\sigma_p^2}{\alpha^2} - \frac{\left(\mu_p + \frac{v_0^T A v_1}{v_1^T A v_1}\right)^2}{\beta^2} = 1$$

(This is the graph of a hyperbola.)

$\alpha^2 > 0$ by Cauchy-Schwarz

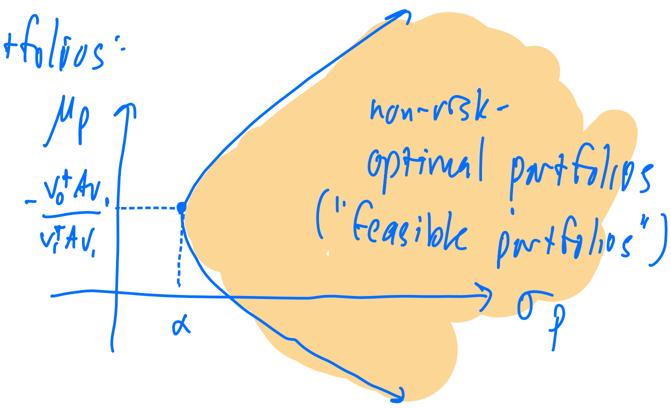
$$\alpha^2 = v_0^T A v_0 - \frac{(v_0^T A v_1)^2}{v_1^T A v_1}$$

$$\beta^2 = \frac{v_0^T A v_0}{v_1^T A v_1} - \left(\frac{v_0^T A v_1}{v_1^T A v_1}\right)^2$$

Vertex: $\mu_p = \frac{-v_0^T A v_1}{v_1^T A v_1}$

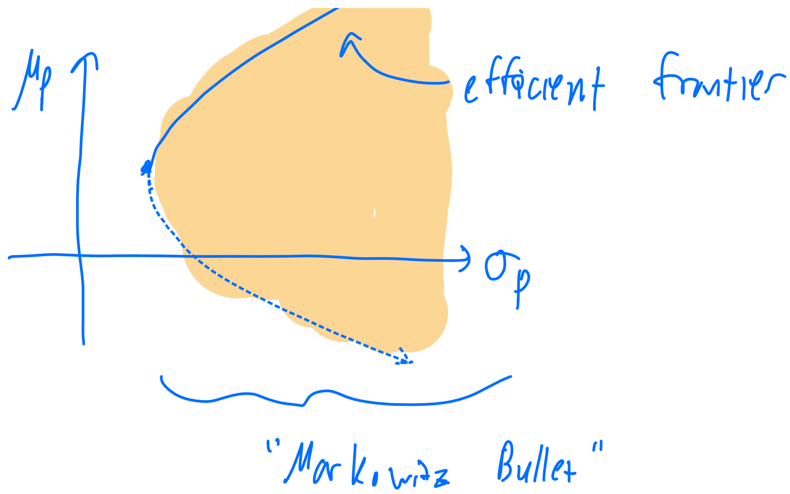
$\sigma_p = \alpha > 0$

Risk-optimal portfolios:



There is a global variance-minimizing portfolio @ the vertex, i.e. at $\mu_p = \frac{-v_0^T A v_1}{v_1^T A v_1} =: \mu_0$

There is an efficient frontier: set of risk-optimal portfolios whose μ_p is maximized for a fixed σ_p.


μ_p
σ_p
efficient frontier
"Markowitz Bullet"

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*

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

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.

Portfolios without short selling

Efficient portfolios can correspond to substantially leveraged portfolios.

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

Portfolios without short selling

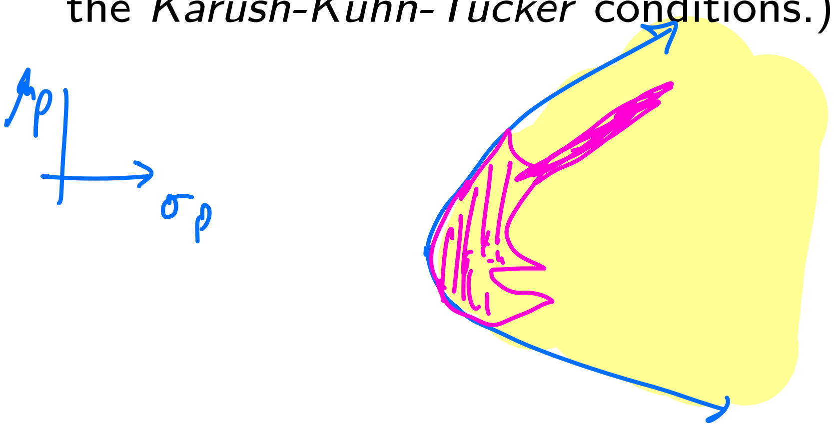
Efficient portfolios can correspond to substantially leveraged portfolios.

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

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

This is an <u>in</u>equality-constrained optimization problem.

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



Portfolios without short selling

Efficient portfolios can correspond to substantially leveraged portfolios.

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

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

This is an <u>in</u>equality-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.