Portfolios that Contain Risky Assets
9.1. Kelly Objectives for Markowitz Portfolios

C. David Levermore

University of Maryland, College Park, MD

Math 420: *Mathematical Modeling*
April 15, 2022 version

Portfolios that Contain Risky Assets
Part II: Probabilistic Models

6. Independent, Identically-Distributed Models for Assets
7. Assessing Independent, Identically-Distributed Models
8. Independent, Identically-Distributed Models for Portfolios
9. Kelly Objectives for Portfolio Models
10. Cautious Objectives for Portfolio Models

Portfolios that Contain Risky Assets Part II: Probabilistic Models

9. Kelly Objectives for Portfolio Models

9.1. Kelly Objectives for Markowitz Portfolios

9.2. Mean-Variance Estimators for Kelly Objectives

9.3. Optimization of Mean-Variance Objectives

Kelly Objectives for Markowitz Portfolios

1. Introduction
2. Kelly for IID Models
3. Sample Estimators
4. One Risk-Free Rate Model
5. Two Risk-Free Rate Model

Introduction

Consider portfolios built from N risky assets and possibly some risk-free assets. If the risky assets are modeled by an IID process then the law of large numbers suggests that the optimal portfolio is selected by the *Kelly criterion*, which is to maximize the expected growth rate over a set of portfolios.

We saw in the context of a simple game that this *Kelly strategy* was not optimal when faced with imperfect knowledge about the game. This led to a discussion of how *fractional Kelly strategies* might compensate for such uncertainty about the game.

Here we begin to apply these ideas to Markowitz portfolios. More specifically, here we will set aside our concerns about uncertainty and develop the Kelly strategy as if our estimators are exact. Later we will show how fractional Kelly strategies emerge when addressing uncertainty.

Kelly for IID Models (Introduction)

We start by reviewing the form the Kelly strategy takes in the setting of an IID model for the returns of risky assets. More specifically, we review the functional form that growth rate means take for Markowitz portfolios. It is these functions that the Kelly strategy seeks to maximize.

An IID model for the returns of the N risky assets draws a sample $\{\mathbf{R}_d\}_{d=1}^{D}$ from a probability density $q(\mathbf{R})$ over $(-1,\infty)^N$. These random variables have mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Xi}$ given by

$$\begin{aligned} \boldsymbol{\mu} &= \mathrm{Ex}(\mathbf{R}) = \int \mathbf{R}\, q(\mathbf{R})\, \mathrm{d}\mathbf{R}\,, \\ \boldsymbol{\Xi} &= \mathrm{Vr}(\mathbf{R}) = \mathrm{Ex}\Big((\mathbf{R}-\boldsymbol{\mu})\,(\mathbf{R}-\boldsymbol{\mu})^{\mathrm{T}} \Big) \\ &= \int (\mathbf{R}-\boldsymbol{\mu})\,(\mathbf{R}-\boldsymbol{\mu})^{\mathrm{T}}\, q(\mathbf{R})\, \mathrm{d}\mathbf{R}\,. \end{aligned} \tag{2.1}$$

Kelly for IID Models (Returns)

Given the IID sample $\{\mathbf{R}_d\}_{d=1}^{D}$, a Markowitz portfolio with risky asset allocation $\mathbf{f}$ has the IID return sample $\{R_d\}_{d=1}^{D}$ given by

$$R_d = r_{\rm rf} + \mathbf{R}_d^{\rm T}\mathbf{f} \qquad \text{for every } d \in \{1, 2, \cdots, D\}\,, \tag{2.2a}$$

where the *risk-free return* $r_{\rm rf}$ is given by

$$r_{\rm rf} = \begin{cases} 0 & \text{when } \mathbf{f} \in \mathcal{M}\,, \\ \mu_{\rm rf} f^{\rm rf} & \text{when } (\mathbf{f}, f^{\rm rf}) \in \mathcal{M}_1\,, \\ \mu_{\rm si} f^{\rm si} + \mu_{\rm cl} f^{\rm cl} & \text{when } (\mathbf{f}, f^{\rm si}, f^{\rm cl}) \in \mathcal{M}_2\,. \end{cases} \tag{2.2b}$$

These random variables have mean μ and variance ξ given by

$$\mu = \mathrm{Ex}(R) = r_{\rm rf} + \boldsymbol{\mu}^{\rm T}\mathbf{f}\,, \qquad \xi = \mathrm{Vr}(R) = \mathbf{f}^{\rm T}\boldsymbol{\Xi}\,\mathbf{f}\,. \tag{2.3}$$

Kelly for IID Models (Growth Rates)

The portfolio is said to be *solvent* over the IID return sample $\{R_d\}_{d=1}^{D}$ if

$$1 + R_d > 0 \qquad \text{for every } d \in \{1, 2, \cdots, D\}\,. \tag{2.4}$$

Every solvent portfolio has the IID growth rate sample $\{X_d\}_{d=1}^{D}$ given by

$$X_d = \log(1 + R_d) \qquad \text{for every } d \in \{1, 2, \cdots, D\}\,. \tag{2.5}$$

These random variables have mean γ and variance θ given by

$$\begin{aligned}
\gamma &= \mathrm{Ex}(X) = \mathrm{Ex}\big(\log(1 + R)\big) = \mathrm{Ex}\Big(\log\Big(1 + r_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big)\Big)\,, \\
\theta &= \mathrm{Vr}(X) = \mathrm{Ex}\Big((X - \gamma)^2\Big) \\
&= \mathrm{Ex}\Big(\big(\log(1 + R) - \gamma\big)^2\Big) = \mathrm{Ex}\bigg(\Big(\log\Big(1 + r_{\mathrm{rf}} + \mathbf{R}^{\mathrm{T}}\mathbf{f}\Big) - \gamma\Big)^2\bigg)\,.
\end{aligned} \tag{2.6}$$

These cannot be expressed exactly in terms of μ and ξ.

Kelly for IID Models (Questions)

The Kelly strategy is to select the portfolio that maximizes the growth rate mean γ given by (2.6) over a set Π of portfolio allocations. This strategy raises several questions.

1. Does such a maximum exist over the set of portfolio allocations being considered? If such a maximum exists, is there a unique maximizer?
2. How can an estimator $\hat{\gamma}$ for γ be built from a sample $\{\mathbf{R}_d\}_{d=1}^{\infty}$?
3. How should the set of portfolio allocations being considered depend upon the estimator $\hat{\gamma}$?
4. Does the estimator $\hat{\gamma}$ have a unique maximizer over the set of portfolio allocations being considered?
5. How close is the maximizer of $\hat{\gamma}$ to that of γ?

Kelly for IID Models (Plan)

These questions will be addressed in upcoming sets of slides.

- The existence and uniqueness of maximizers raised by questions 1 and 4 will be addressed by showing that the continuous objective functions are strictly concave over the convex set Π and cannot have a maximum outside of a compact subset of Π.
- Several estimators $\hat{\gamma}$ will be built and analyzed. Appropriate sets of portfolio allocations will be identified for each, thereby addressing questions 2 and 3.
- Explicit maximizers for some of these estimators will be constructed for certain sets of Markowitz portfolios, some without risk-free assets, some with the one-rate model, and others with the two-rate model. Comparing these will begin to address question 5.

Sample Estimators (Means and Variances)

When an IID model is used for the returns of risky assets, we treat a return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ as if, or almost as if it was an IID sample drawn from an unknown probability density. We then choose positive weights $\{w_d\}_{d=1}^{D}$ that sum to 1 and set

$$\mathbf{m} = \sum_{d=1}^{D} w_d \, \mathbf{r}(d) \,, \qquad \mathbf{V} = \sum_{d=1}^{D} w_d \left(\mathbf{r}(d) - \mathbf{m}\right) \left(\mathbf{r}(d) - \mathbf{m}\right)^{\mathrm{T}} . \tag{3.7}$$

The return mean $\boldsymbol{\mu}$ and return variance $\boldsymbol{\Xi}$ given by (2.1) for the underlying probability density then have the unbiased estimators

$$\hat{\boldsymbol{\mu}} = \mathbf{m} \,, \qquad \hat{\boldsymbol{\Xi}} = \frac{1}{1 - \bar{w}} \mathbf{V} \qquad \text{where} \qquad \bar{w} = \sum_{d=1}^{D} w_d^2 \,. \tag{3.8}$$

If we have great confidence in the validity if the IID model then we would choose uniform weights. We might choose nonuniform weights if we think that older data is less informative than more recent data.

Sample Estimators (Markowitz Portfolios)

The Markowitz portfolio with the risk-free return $r_{\rm rf}$ and the risky asset allocation $\mathbf{f}$ has the return history $\{r(d)\}_{d=1}^{D}$ given by

$$r(d) = r_{\rm rf} + \mathbf{r}(d)^{\rm T}\mathbf{f}\,, \tag{3.9}$$

where $r_{\rm rf}$ is given in terms of the risk-free asset allocations by (2.2b).

We see from the unbiased estimators for $\boldsymbol{\mu}$ and $\boldsymbol{\Xi}$ given by (3.8) that the return mean μ and return variance ξ given by (2.3) for this portfolio have the unbiased estimators

$$\hat{\mu} = r_{\rm rf} + \mathbf{m}^{\rm T}\mathbf{f}\,, \qquad \hat{\xi} = \frac{1}{1-\bar{w}}\,\mathbf{f}^{\rm T}\mathbf{V}\,\mathbf{f}\,. \tag{3.10}$$

Sample Estimators (Growth Rate Means)

It is evident from the definition of a solvent portfolio given by (2.4) that the Markowitz portfolio with the risk-free return r_{rf} and the risky asset allocation $\mathbf{f}$ is solvent with respect to the return histroy $\{\mathbf{r}(d)\}_{d=1}^{D}$ provided

$$1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0 \qquad \text{for every } d \in \{1, 2, \cdots, D\}\,. \tag{3.11}$$

The growth rate mean γ given by (2.6) for such a portfolio then has the unbiased estimator

$$\hat{\gamma} = \sum_{d=1}^{D} w_d \log\left(1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\right)\,. \tag{3.12}$$

Clearly the solvency condition (3.11) is both necessary and sufficient for $\hat{\gamma}$ given by (3.12) to be defined.

Sample Estimators (Solvent Allocations)

More specifically, given the return history $\{\mathbf{r}(d)\}_{d=1}^{D}$, we see from (2.2) and (3.11) that associated with the sets of Markowitz allocations

$$\begin{aligned} \mathcal{M} &= \left\{ \mathbf{f} \in \mathbb{R}^N \,:\, \mathbf{1}^{\mathrm{T}}\mathbf{f} = 1 \right\}, \\ \mathcal{M}_1 &= \left\{ (\mathbf{f}, f^{\mathrm{rf}}) \in \mathbb{R}^{N+1} \,:\, \mathbf{1}^{\mathrm{T}}\mathbf{f} + f^{\mathrm{rf}} = 1 \right\}, \\ \mathcal{M}_2 &= \left\{ (\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \mathbb{R}^{N+2} \,:\, \mathbf{1}^{\mathrm{T}}\mathbf{f} + f^{\mathrm{si}} + f^{\mathrm{cl}} = 1\,,\ f^{\mathrm{si}} \geq 0\,,\ f^{\mathrm{cl}} \leq 0 \right\}, \end{aligned} \tag{3.13}$$

are the sets of solvent Markowitz allocations

$$\begin{aligned} \Omega &= \left\{ \mathbf{f} \in \mathcal{M} \,:\, 1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0 \ \ \forall d \right\}, \\ \Omega_1 &= \left\{ (\mathbf{f}, f^{\mathrm{rf}}) \in \mathcal{M}_1 \,:\, 1 + \mu_{\mathrm{rf}} f^{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0 \ \ \forall d \right\}, \\ \Omega_2 &= \left\{ (\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \mathcal{M}_2 \,:\, 1 + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0 \ \ \forall d \right\}, \end{aligned} \tag{3.14}$$

Sample Estimators (Growth Rate Means)

We then see from (3.12) that for every $\mathbf{f} \in \Omega$ we have

$$\hat{\gamma}(\mathbf{f}) = \sum_{d=1}^{D} w_d \, \log\Big(1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big) \; , \tag{3.15a}$$

for every $(\mathbf{f}, f^{\mathrm{rf}}) \in \Omega_1$ we have

$$\hat{\gamma}(\mathbf{f}, f^{\mathrm{rf}}) = \sum_{d=1}^{D} w_d \, \log\Big(1 + \mu_{\mathrm{rf}} f^{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big) \; , \tag{3.15b}$$

and for every $(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \Omega_2$ we have

$$\hat{\gamma}(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) = \sum_{d=1}^{D} w_d \, \log\Big(1 + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big) \; . \tag{3.15c}$$

Sample Estimators (Convexity)

We now give some facts about the sets Ω, Ω_1 and Ω_2 defined by (3.14) and the functions $\hat{\gamma}$ defined over them by (3.15).

Fact 1. The sets Ω, Ω_1 and Ω_2 defined by (3.14) are open, convex subsets of the sets $\mathcal{M}$, $\mathcal{M}_1$ and $\mathcal{M}_2$ given by (3.13) respectively.

Proof. Each solvency constraint in the definitions (3.14) of Ω, Ω_1 and Ω_2 has the general form

$$1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0 \qquad \text{for some } d\,,$$

which describes an open convex subset of $\mathcal{M}$, $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively. Beacuse Ω, Ω_1 and Ω_2 are the intersections of their constraints, they are open convex subsets of $\mathcal{M}$, $\mathcal{M}_1$ and $\mathcal{M}_2$ respectively. $\square$

Sample Estimators (Jensen Inequality Bound)

The growth rate mean estimators given by (3.15) derive from the general form (3.12), which is

$$\hat{\gamma} = \sum_{d=1}^{D} w_d \log\Big(1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big) .$$

Because $\log(1 + r)$ is a concave function of r over $(-1, \infty)$, we can apply the Jensen inequality to this general form to obtain the following.

Fact 2. The growth rate mean estimator $\hat{\gamma}$ given by (3.12) and the return mean estimator given by (3.10) are related by the *Jensen inequality bound*

$$\hat{\gamma} \leq \log(1 + \hat{\mu}) . \tag{3.16}$$

We will prove this bound after reviewing the Jensen inequality.

Sample Estimators (Jensen Inequality)

Jensen Inequality. Let $g(z)$ be a convex (concave) function over an interval $[a, b]$. Let the points $\{z_d\}_{d=1}^{D}$ lie within $[a, b]$. Let $\{w_d\}_{d=1}^{D}$ be nonnegative weights that sum to one. Then

$$g(\bar{z}) \leq \overline{g(z)} \qquad \left(\overline{g(z)} \leq g(\bar{z})\right), \tag{3.17a}$$

where

$$\bar{z} = \sum_{d=1}^{D} z_d \, w_d \,, \qquad \overline{g(z)} = \sum_{d=1}^{D} g(z_d) \, w_d \,. \tag{3.17b}$$

Remark. There is an integral version of the Jensen inequality that we do not give here because we do not need it.

Sample Estimators (Jensen Proof)

Proof of the Jensen Inequality. We consider the case when $g(z)$ is convex and differentiable over $[a,b]$. Then for every $\bar{z} \in [a,b]$ we have

$$g(z) \geq g(\bar{z}) + g'(\bar{z})(z - \bar{z}) \qquad \text{for every } z \in [a,b]\,.$$

This inequality simply says that the tangent line to the graph of g at $\bar{z}$ lies below the graph of g over $[a,b]$. Let $\bar{z}$ be given by (3.17b). By then setting $z = z_d$ in the above inequality, multiplying both sides by w_d, and summing over d we obtain

$$\begin{aligned}\overline{g(z)} = \sum_{d=1}^{D} g(z_d)\, w_d &\geq \sum_{d=1}^{D} \Big(g(\bar{z}) + g'(\bar{z})(z_d - \bar{z}) \Big)\, w_d \\ &= g(\bar{z}) \sum_{d=1}^{D} w_d + g'(\bar{z}) \left(\sum_{d=1}^{D} (z_d - \bar{z})\, w_d \right) = g(\bar{z})\,.\end{aligned}$$

This proves the Jensen inequality (3.17a). $\square$

Sample Estimators (**Fact 2** Proof)

Proof of Fact 2. Let the Markowitz portfolio with the risk-free return r_{rf} and the risky asset allocation $\mathbf{f}$ be solvent. This implies by (3.11) that the set of points $\{r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\}_{d=1}^{D}$ lies within an interval $[a, b] \subset (-1, \infty)$. Because $\log(1 + r)$ is a concave function of r over $(-1, \infty)$, definition (3.12) of $\hat{\gamma}$, the Jensen inequality (3.17) applied to $g(z) = \log(1 + z)$ with $z_d = r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}$, and definition (3.10) of $\hat{\mu}$ yield

$$\begin{aligned}
\hat{\gamma} &= \sum_{d=1}^{D} w_d \, \log\Big(1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big) \\
&\leq \log\left(1 + \sum_{d=1}^{D} w_d \, \Big(1 + r_{\mathrm{rf}} + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\Big)\right) \\
&= \log(1 + \hat{\mu}) \, .
\end{aligned}$$

This establishes the bound (3.16), whereby **Fact 2** is proved. □

Sample Estimators (Continuity and Concavity)

Fact 3. The $\hat{\gamma}$ given by (3.15) are continuous, concave functions over the convex sets Ω, Ω_1 and Ω_2 respectively.

Proof. Because $\log(y)$ is continuous and concave over $y > 0$, we see that the $\hat{\gamma}$ given by (3.15) are sums of continuous, concave functions over the convex sets Ω, Ω_1 and Ω_2 respectively. Therefore the $\hat{\gamma}$ given by (3.15) are continuous, concave functions over the convex sets Ω, Ω_1 and Ω_2 respectively. □

Later we will show the following.

Fact 4. If $\mathbf{V}$ is positive definite then the $\hat{\gamma}$ given by (3.15) are strictly concave functions over Ω, Ω_1 and Ω_2 respectively.

Sample Estimators (Allocations Π)

The set Π of portfolio allocations considered can be a convex subset of either Ω, Ω_1 or Ω_2 depending on how risk-free assets are modeled.

- For portfolios with no risk-free assets we consider some convex set $\Pi \subset \Omega$. In this case the Kelly objective $\hat{\gamma}$ is given by (3.15a).
- For portfolios with risk-free assets then for the one rate model we consider some convex set $\Pi \subset \Omega_1$. In this case the Kelly objective $\hat{\gamma}$ is given by (3.15b).
- For portfolios with risk-free assets then for the two rate model we consider some convex set $\Pi \subset \Omega_2$. In this case the Kelly objective $\hat{\gamma}$ is given by (3.15c).

Sample Estimators (Maximizer)

We have the following general considerations regarding the existence and uniqueness of a maximizer for the Kelly objective $\hat{\gamma}$ over the convex set Π.

- By **Fact 4** if $\mathbf{V}$ is positive definite then $\hat{\gamma}$ is strictly concave over the convex set Π, whereby if $\hat{\gamma}$ has a maximizer over Π then it is unique.
- If Π is a compact set then the Extreme-Value Theorem insures the existence of a maximizer because $\hat{\gamma}$ is continuous over Π by **Fact 3**. For example, Π can be a set of long allocations (like Λ or Λ_1), a set of limited-leverage allocations (like Π^ℓ, Π_1^ℓ or Π_2^ℓ with ℓ small enough that the portfolios are solvent), or some other compact, convex set.
- If Π is not a compact set then some analysis needs to be done to show that a maximizer exists.
  - If Π is not closed then we must check the behavior of $\hat{\gamma}$ near boundary points of Π that are not in Π.
  - If Π is not bounded then we must check the behavior of $\hat{\gamma}$ for unbounded sequences in Π.

Sample Estimators (Example)

Example. For many return historys $\{\mathbf{r}(d)\}_{d=1}^{D}$ the set Ω will be bounded. Suppose that we are in that case and that $\Pi = \Omega$. The definition of Ω given in (3.14) implies that if $\{\mathbf{f}_n\}_{n=1}^{\infty}$ is a sequence in Ω that approaches a boundary point of Ω then there will be at least one d for which

$$1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}_n \searrow 0 \qquad \text{as } n \to \infty \,.$$

For every such d we have

$$\log\Big(1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}_n\Big) \to -\infty \qquad \text{as } n \to \infty \,.$$

It follows from (3.15a) that $\hat{\gamma}(\mathbf{f}_n) \to -\infty$ as $n \to \infty$ for this sequence. Because $\hat{\gamma}(\mathbf{f})$ is continuous over the bounded set Ω and goes to $-\infty$ as $\mathbf{f}$ approaches the boundary of Ω, it follows that $\hat{\gamma}(\mathbf{f})$ has a maximizer in Ω. Because $\hat{\gamma}(\mathbf{f})$ is strictly concave over Ω, this maximizer is unique.

One Risk-Free Rate Model (Introduction)

Here we will prove **Fact 4** for the one risk-free rate model. As a bonus, we will also prove it for portfolios with no risk-free assets. Rather than working in $\mathcal{M}_1$ given by (3.13), we will use the constraint $\mathbf{1}^{\mathrm{T}}\mathbf{f} + f^{\mathrm{rf}} = 1$ to eliminate f^{rf} and work in $\mathcal{M}_+ = \mathbb{R}^N$. The set of solvent allocations is

$$\Omega_+ = \left\{ \mathbf{f} \in \mathcal{M}_+ \,:\, 1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f} > 0 \ \ \forall d \right\}. \tag{4.18}$$

The growth rate mean estimator (3.15b) then becomes

$$\hat{\gamma}(\mathbf{f}) = \sum_{d=1}^{D} w_d \log\Big(1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f}\Big) . \tag{4.19}$$

The case of portfolios with no risk-free assets is included because

$$\Omega = \left\{ \mathbf{f} \in \Omega_+ \,:\, \mathbf{1}^{\mathrm{T}}\mathbf{f} = 1 \right\},$$

and $\hat{\gamma}(\mathbf{f})$ given by (4.19) agrees with $\hat{\gamma}(\mathbf{f})$ given by (3.15a) on this set.

One Risk-Free Rate Model (Derivatives)

Because $\hat{\gamma}(\mathbf{f})$ given by (4.19) is

$$\hat{\gamma}(\mathbf{f}) = \sum_{d=1}^{D} w_d \, \log\Big(1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f}\Big) \,,$$

we see that it is an infinitely differentiable function of $\mathbf{f}$ over Ω_+ with

$$\begin{aligned} \nabla_{\mathbf{f}}\, \hat{\gamma}(\mathbf{f}) &= \sum_{d=1}^{D} w_d \, \frac{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}}{1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f}} \,, \\ \nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f}) &= -\sum_{d=1}^{D} w_d \, \frac{(\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})\,(\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}}{(1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f})^2} \,. \end{aligned} \tag{4.20}$$

One Risk-Free Rate Model (Hessian Matrix)

The Hessian matrix $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ has the following properties.

Fact 5. Let $\mathbf{f} \in \Omega_+$ be arbitrary.

- The matrix $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is nonpositive definite.
- The matrix $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is negative definite if and only if

$$\text{the vectors } \{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}\}_{d=1}^{D} \text{ span } \mathbb{R}^N . \tag{4.21}$$

Remark. **Fact 5** implies that $\hat{\gamma}(\mathbf{f})$ is concave over Ω_+, which was already proven in **Fact 3**. Moreover, it implies that $\hat{\gamma}(\mathbf{f})$ is strictly concave over Ω_+ when the vectors $\{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}\}_{d=1}^{D}$ span $\mathbb{R}^N$. Because this condition holds when $\mathbf{V}$ is positive definite, **Fact 4** will follow from **Fact 5**.

One Risk-Free Rate Model (**Fact 5** Proof)

Proof of Fact 5. Let $\mathbf{f} \in \Omega_+$ be arbitrary. Then for every $\mathbf{y} \in \mathbb{R}^N$ we see from (4.20) that

$$\mathbf{y}^{\mathrm{T}} \nabla_{\mathbf{f}}^2 \, \hat{\gamma}(\mathbf{f}) \, \mathbf{y} = - \sum_{d=1}^{D} w_d \, \frac{\mathbf{y}^{\mathrm{T}} (\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1}) \, (\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1})^{\mathrm{T}} \mathbf{y}}{(1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1})^{\mathrm{T}} \mathbf{f})^2}$$
$$= - \sum_{d=1}^{D} w_d \, \frac{\left((\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1})^{\mathrm{T}} \mathbf{y} \right)^2}{(1 + \mu_{\mathrm{rf}} + (\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1})^{\mathrm{T}} \mathbf{f})^2} \leq 0 \, .$$

Therefore $\nabla_{\mathbf{f}}^2 \, \hat{\gamma}(\mathbf{f})$ is nonpositive definite. Moreover, we see that

$$\mathbf{y}^{\mathrm{T}} \nabla_{\mathbf{f}}^2 \, \hat{\gamma}(\mathbf{f}) \, \mathbf{y} = 0 \quad \Longleftrightarrow \quad \left(\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1}\right)^{\mathrm{T}} \mathbf{y} = 0 \;\; \forall d \, .$$

Therefore $\nabla_{\mathbf{f}}^2 \, \hat{\gamma}(\mathbf{f})$ is negative definite if and only if

$$\left(\mathbf{r}(d) - \mu_{\mathrm{rf}} \mathbf{1}\right)^{\mathrm{T}} \mathbf{y} = 0 \;\; \forall d \quad \Longrightarrow \quad \mathbf{y} = \mathbf{0} \, . \tag{4.22}$$

One Risk-Free Rate Model (**Fact 5** Proof)

To complete the proof of **Fact 5** we need to show that implication (4.22) holds if and only if the spanning condition (4.21) is met.

First, suppose that implication (4.22) does not hold. Then there exists $\mathbf{y} \in \mathbb{R}^N$ such that

$$\mathbf{y} \neq \mathbf{0} \qquad \text{and} \qquad (\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{y} = 0 \quad \forall d\,. \tag{4.23}$$

This says that the vectors $\{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}\}_{d=1}^{D}$ all lie in the linear subspace orthogonal (normal) to $\mathbf{y}$. Therefore they do not span $\mathbb{R}^N$, so the spanning condition (4.21) is not met.

Conversely, suppose that the spanning condition (4.21) is not met. Then $\mathrm{Span}\Big\{\{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}\}_{d=1}^{D}\Big\}$ is a proper linear subspace of $\mathbb{R}^N$, so there exists a nonzero $\mathbf{y} \in \mathbb{R}^N$ orthogonal to it. But this says that $\mathbf{y}$ satisfies (4.23), so implication (4.22) does not hold.

The proof of **Fact 5** is now complete.

One Risk-Free Rate Model (**Fact 4** Proof)

Proof of Fact 4 for the One Risk-Free Rate Model. Recall that we assumed that the covariance matrix $\mathbf{V}$ is positive definite. Recall too that this is equivalent to assuming that

$$\text{the set } \{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D} \text{ spans } \mathbb{R}^N .$$

But this condition implies that

$$\text{the set } \{\mathbf{r}(d) - \mu_{\mathrm{rf}}\mathbf{1}\}_{d=1}^{D} \text{ spans } \mathbb{R}^N ,$$

so by **Fact 5** it implies that $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is negative definite for every $\mathbf{f} \in \Omega_+$. Therefore $\hat{\gamma}(\mathbf{f})$ given by (4.19) is a strictly concave function over Ω_+. □

One Risk-Free Rate Model (Unique Maximizer)

Fact 6. Let $\mathbf{V}$ be positive definite. Let $\Pi \subset \Omega_+$ be convex. If $\hat{\gamma}(\mathbf{f})$ has a maximum over Π then it has a unique maximizer in Π.

Proof. Because $\mathbf{V}$ is positive definite, **Fact 4** implies that $\hat{\gamma}(\mathbf{f})$ is a strictly concave function of $\mathbf{f}$ over the convex set Π. Suppose that $\hat{\gamma}(\mathbf{f})$ has a maximum $\hat{\gamma}_{\mathrm{mx}}$ over Π, and that $\mathbf{f}_0$ and $\mathbf{f}_1 \in \Pi$ are maximizers with $\mathbf{f}_0 \neq \mathbf{f}_1$. For every $t \in (0,1)$ define $\mathbf{f}_t = (1-t)\,\mathbf{f}_0 + t\,\mathbf{f}_1$. Then for every $t \in (0,1)$ the convexity of Π implies that $\mathbf{f}_t \in \Pi$ while the strict concavity of $\hat{\gamma}(\mathbf{f})$ over Π implies that

$$\begin{aligned}\hat{\gamma}(\mathbf{f}_t) &> (1-t)\,\hat{\gamma}(\mathbf{f}_0) + t\,\hat{\gamma}(\mathbf{f}_1) \\ &= (1-t)\,\hat{\gamma}_{\mathrm{mx}} + t\,\hat{\gamma}_{\mathrm{mx}} = \hat{\gamma}_{\mathrm{mx}}\,.\end{aligned}$$

But this contradicts the fact that $\hat{\gamma}_{\mathrm{mx}}$ is the maximum of $\hat{\gamma}(\mathbf{f})$ over Π. Therefore at most one maximizer can exist. □

Two Risk-Free Rate Model (Introduction)

Here we will prove **Fact 4** for the two risk-free rate model. Recall from (3.13), (3.14) and (3.15c) that for this model we have

$$\mathcal{M}_2 = \Big\{(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \mathbb{R}^{N+2} \; : \; \mathbf{1}^{\mathrm{T}}\mathbf{f} + f^{\mathrm{si}} + f^{\mathrm{cl}} = 1\,,\; f^{\mathrm{si}} \geq 0\,,\; f^{\mathrm{cl}} \leq 0 \Big\}\,, \tag{5.24a}$$

$$\Omega_2 = \Big\{(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \mathcal{M}_2 \; : \; 1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}} > 0 \;\; \forall d \Big\}\,, \tag{5.24b}$$

$$\hat{\gamma}(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) = \sum_{d=1}^{D} w_d \log\Big(1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}}\Big)\,. \tag{5.24c}$$

The natural domain of $\hat{\gamma}(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}})$ is an open set in $\mathbb{R}^{N+2}$ containing Ω_2.

Two Risk-Free Rate Model (**Fact 4** Proof)

Proof of Fact 4. The estimator $\hat{\gamma}(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}})$ is infinitely differentiable over its natrual domain. Let $\mathrm{Hess}(\hat{\gamma})(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}})$ denote the Hessian of $\hat{\gamma}$ at some $(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \Omega_2$, so that

$$\mathrm{Hess}(\hat{\gamma}) = \begin{pmatrix} \nabla_{\mathbf{f}}^2 \hat{\gamma} & \nabla_{\mathbf{f}} \partial_{f^{\mathrm{si}}} \hat{\gamma} & \nabla_{\mathbf{f}} \partial_{f^{\mathrm{cl}}} \hat{\gamma} \\ \partial_{f^{\mathrm{si}}} \nabla_{\mathbf{f}} \hat{\gamma}^{\mathrm{T}} & \partial_{f^{\mathrm{si}}}^2 \hat{\gamma} & \partial_{f^{\mathrm{si}}} \partial_{f^{\mathrm{cl}}} \hat{\gamma} \\ \partial_{f^{\mathrm{cl}}} \nabla_{\mathbf{f}} \hat{\gamma}^{\mathrm{T}} & \partial_{f^{\mathrm{cl}}} \partial_{f^{\mathrm{si}}} \hat{\gamma} & \partial_{f^{\mathrm{cl}}}^2 \hat{\gamma} \end{pmatrix} .$$

Because of the equality constraint in definition (5.24a) of $\mathcal{M}_2$, and because $\Omega_2 \subset \mathcal{M}_2$, tangent vectors of Ω_2 must satisfy the constraint

$$\mathbf{1}^{\mathrm{T}} \mathbf{y} + y^{\mathrm{si}} + y^{\mathrm{cl}} = 0 . \tag{5.25}$$

The tangent space of Ω_2 is comprised of all such vectors.

Two Risk-Free Rate Model (**Fact 4** Proof)

It can be shown for any $(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \Omega_2$ and any $(\mathbf{y}, y^{\mathrm{si}}, y^{\mathrm{cl}}) \in \mathbb{R}^{N+2}$ that

$$
\begin{pmatrix} \mathbf{y}^{\mathrm{T}} & y^{\mathrm{si}} & y^{\mathrm{cl}} \end{pmatrix} \mathrm{Hess}(\hat{\gamma})(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \begin{pmatrix} \mathbf{y} \\ y^{\mathrm{si}} \\ y^{\mathrm{cl}} \end{pmatrix}
= - \sum_{d=1}^{D} w_d \, \frac{\left(\mathbf{r}(d)^{\mathrm{T}} \mathbf{y} + \mu_{\mathrm{si}} y^{\mathrm{si}} + \mu_{\mathrm{cl}} y^{\mathrm{cl}} \right)^2}{(1 + \mathbf{r}(d)^{\mathrm{T}} \mathbf{f} + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}})^2} \,.
$$

Because tangent vectors of Ω_2 satisfy (5.25), we see that

- $\mathrm{Hess}(\hat{\gamma})$ is nonpositive definite.
- $\mathrm{Hess}(\hat{\gamma})$ is negative definite over the tangent space of Ω_2 if and only if

$$
\left.\begin{aligned} \mathbf{1}^{\mathrm{T}} \mathbf{y} + y^{\mathrm{si}} + y^{\mathrm{cl}} &= 0 \\ \mathbf{r}(d)^{\mathrm{T}} \mathbf{y} + \mu_{\mathrm{si}} y^{\mathrm{si}} + \mu_{\mathrm{cl}} y^{\mathrm{cl}} &= 0 \quad \forall d \end{aligned}\right\} \implies \begin{pmatrix} \mathbf{y} \\ y^{\mathrm{si}} \\ y^{\mathrm{cl}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 0 \\ 0 \end{pmatrix} . \qquad (5.26)
$$

Two Risk-Free Rate Model (**Fact 4** Proof)

Because

$$\mathbf{r}(d)^{\mathrm{T}}\mathbf{y} + \mu_{\mathrm{si}} y^{\mathrm{si}} + \mu_{\mathrm{cl}} y^{\mathrm{cl}} = 0 \quad \forall d \quad \Longrightarrow \quad \mathbf{m}^{\mathrm{T}}\mathbf{y} + \mu_{\mathrm{si}} y^{\mathrm{si}} + \mu_{\mathrm{cl}} y^{\mathrm{cl}} = 0\,,$$

we see that

$$\left(\mathbf{r}(d) - \mathbf{m}\right)^{\mathrm{T}}\mathbf{y} = 0 \quad \forall d\,.$$

But this implies $\mathbf{V}\mathbf{y} = 0$, which says $\mathbf{y} = \mathbf{0}$ because $\mathbf{V}$ is positive definite. The linear system in (5.26) thereby reduces to the 2×2 system

$$\begin{aligned} y^{\mathrm{si}} + y^{\mathrm{cl}} &= 0\,, \\ \mu_{\mathrm{si}} y^{\mathrm{si}} + \mu_{\mathrm{cl}} y^{\mathrm{cl}} &= 0\,. \end{aligned}$$

Because $\mu_{\mathrm{cl}} > \mu_{\mathrm{si}}$, this system shows that $y^{\mathrm{si}} = y^{\mathrm{cl}} = 0$. Thus, we have proved implication (5.26), which shows that $\mathrm{Hess}(\hat{\gamma})$ is negative definite over the tangent space of Ω_2. Therefore $\hat{\gamma}$ is strictly concave over Ω_2.