Portfolios that Contain Risky Assets 14: Kelly Objectives for Markowitz Portfolios

C. David Levermore

University of Maryland, College Park, MD

Math 420: *Mathematical Modeling*
April 27, 2019 version

Portfolios that Contain Risky Assets
Part II: Stochastic Models

Kelly Objectives for Markowitz Portfolios

Introduction

We now apply the Kelly criterion to classes of Markowitz portfolios. Given a daily return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ on N risky assets, a daily return μ_{si} on a safe investment, and a daily return μ_{cl} on a credit line, the Markowitz portfolio with allocation $\mathbf{f}$ in risky assets has the daily return history $\{r(d,\mathbf{f})\}_{d=1}^{D}$, where

$$r(d,\mathbf{f}) = \mu_{\mathrm{rf}}(\mathbf{f})(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}) + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\,, \tag{1.1}$$

with

$$\mu_{\mathrm{rf}}(\mathbf{f}) = \begin{cases} \mu_{\mathrm{si}} & \text{if } \mathbf{1}^{\mathrm{T}}\mathbf{f} \leq 1\,, \\ \mu_{\mathrm{cl}} & \text{if } \mathbf{1}^{\mathrm{T}}\mathbf{f} > 1\,. \end{cases} \tag{1.2}$$

The *one risk-free rate model* for risk-free assets assumes $0 < \mu_{\mathrm{si}} = \mu_{\mathrm{cl}}$. The *two risk-free rate model* for risk-free assets assumes $0 < \mu_{\mathrm{si}} < \mu_{\mathrm{cl}}$. Portfolios *without risk-free assets* satisfy the constraint $\mathbf{1}^{\mathrm{T}}\mathbf{f} = 1$.

Introduction

We will consider only classes of *solvent Markowitz portfolios*. This means that we require $\mathbf{f} \in \Omega^+$ where

$$\Omega^+ = \left\{ \mathbf{f} \in \mathbb{R}^N \,:\, 1 + r(d, \mathbf{f}) > 0 \quad \forall d \right\} . \tag{1.3}$$

We can show that *$r(d, \mathbf{f})$ is a concave function of $\mathbf{f}$ over $\mathbb{R}^N$ for every d*. This means that for every d and every $\mathbf{f}_0$, $\mathbf{f}_1 \in \mathbb{R}^N$ we can show that

$$r(d, \mathbf{f}_t) \geq (1 - t)\, r(d, \mathbf{f}_0) + t\, r(d, \mathbf{f}_1) \quad \text{for every } t \in [0, 1] ,$$

where $\mathbf{f}_t = (1 - t)\, \mathbf{f}_0 + t\, \mathbf{f}_1$. This concavity implies that for every $\mathbf{f}_0$, $\mathbf{f}_1 \in \Omega^+$ and every $t \in [0, 1]$ we have

$$\begin{aligned} 1 + r(d, \mathbf{f}_t) &\geq 1 + (1 - t)\, r(d, \mathbf{f}_0) + t\, r(d, \mathbf{f}_1) \\ &= (1 - t)\, \big(1 + r(d, \mathbf{f}_0)\big) + t\, \big(1 + r(d, \mathbf{f}_1)\big) \geq 0 , \end{aligned}$$

whereby $\mathbf{f}_t \in \Omega^+$. Therefore *$\Omega^+$ is a convex set*.

Introduction

The solvent Markowitz portfolio with allocation $\mathbf{f}$ has the growth rate history $\{x(d,\mathbf{f})\}_{d=1}^{D}$ where

$$x(d,\mathbf{f}) = \log\big(1 + r(d,\mathbf{f})\big)\,. \tag{1.4}$$

Notice that the growth rate history is only defined for solvent portfolios.

Because $r(d,\mathbf{f})$ is a concave function over $\mathbf{f} \in \mathbb{R}^N$ for every d while $\log(1+r)$ is an increasing, strictly concave function of r over $r \in (-1,\infty)$, we can show that *$x(d,\mathbf{f})$ is a concave function of $\mathbf{f}$ over Ω^+ for every d*. Indeed, for every $\mathbf{f}_0$, $\mathbf{f}_1 \in \Omega^+$ and every $t \in [0,1]$ we have

$$\begin{aligned}
x(d,\mathbf{f}_t) &= \log\big(1 + r(d,\mathbf{f}_t)\big) \\
&\geq \log\big(1 + (1-t)\,r(d,\mathbf{f}_0) + t\,r(d,\mathbf{f}_1)\big) \\
&\geq (1-t)\,\log\big(1 + r(d,\mathbf{f}_0)\big) + t\,\log\big(1 + r(d,\mathbf{f}_1)\big) \\
&= (1-t)\,x(d,\mathbf{f}_0) + t\,x(d,\mathbf{f}_1)\,.
\end{aligned}$$

Sample Estimators of the Growth Rate Mean

If we use an IID model for the class of solvent Markowitz portfolios then the Kelly criterion says that for maximal long-term growth we should pick $\mathbf{f} \in \Omega^+$ to maximize the growth rate mean $\gamma(\mathbf{f})$. Because we do not know $\gamma(\mathbf{f})$, we might maximize an estimator for $\gamma(\mathbf{f})$. Here we explore sample esitmators of $\gamma(\mathbf{f})$.

Given and allocation $\mathbf{f}$ and weights $\{w_d\}_{d=1}^{D}$ such that

$$w_d > 0 \quad \forall d\,, \qquad \sum_{d=1}^{D} w_d = 1\,, \tag{2.5}$$

the growth rate history $\{x(d,\mathbf{f})\}_{d=1}^{D}$ yields the sample estimator

$$\hat{\gamma}(\mathbf{f}) = \sum_{d=1}^{D} w_d\, x(d,\mathbf{f}) = \sum_{d=1}^{D} w_d \log\big(1 + r(d,\mathbf{f})\big)\,. \tag{2.6}$$

This is clearly defined for every $\mathbf{f} \in \Omega^+$.

Sample Estimators of the Growth Rate Mean

Here are some facts about $\hat{\gamma}(\mathbf{f})$ considered as a function over Ω^+.

Fact 1. $\hat{\gamma}(\mathbf{0}) = \log(1 + \mu_{\mathrm{si}})$.

Fact 2. $\hat{\gamma}(\mathbf{f})$ *is concave over* Ω^+.

Fact 3. *For every* $\mathbf{f} \in \Omega^+$ *we have the bound*

$$\hat{\gamma}(\mathbf{f}) \leq \log\big(1 + \hat{\mu}(\mathbf{f})\big)\,, \tag{2.7}$$

where $\hat{\mu}(\mathbf{f})$ is the sample estimator of the return mean given by

$$\begin{aligned} \hat{\mu}(\mathbf{f}) = \sum_{d=1}^{D} w_d\, r(d,\mathbf{f}) &= \mu_{\mathrm{rf}}(\mathbf{f})\big(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}\big) + \sum_{d=1}^{D} w_d\, \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} \\ &= \mu_{\mathrm{rf}}(\mathbf{f})\big(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}\big) + \mathbf{m}^{\mathrm{T}}\mathbf{f}\,. \end{aligned} \tag{2.8}$$

Remark. **Fact 1** shows that bound (2.7) is an equality when $\mathbf{f} = \mathbf{0}$.

Sample Estimators of the Growth Rate Mean

Proof of Fact 1. Definitions (1.1) and (1.2) of $r(d,\mathbf{f})$ and $\mu_{\mathrm{rf}}(\mathbf{f})$ respectively show that

$$r(d,\mathbf{0}) = \mu_{\mathrm{rf}}(\mathbf{0})(1 - \mathbf{1}^{\mathrm{T}}\mathbf{0}) + \mathbf{r}(d)^{\mathrm{T}}\mathbf{0} = \mu_{\mathrm{rf}}(\mathbf{0}) = \mu_{\mathrm{si}}\,.$$

Then definition (2.6) of $\hat{\gamma}(\mathbf{f})$ yields

$$\begin{aligned}\hat{\gamma}(\mathbf{0}) &= \sum_{d=1}^{D} w_d \, \log\big(1 + r(d,\mathbf{0})\big) \\ &= \sum_{d=1}^{D} w_d \, \log(1 + \mu_{\mathrm{si}}) \\ &= \log(1 + \mu_{\mathrm{si}})\,.\end{aligned}$$

Therefore we have proved **Fact 1**. □

Sample Estimators of the Growth Rate Mean

Proof of Fact 2. Because $x(d, \mathbf{f})$ is a concave function of $\mathbf{f}$ over Ω^+ for every d, and because definition (2.6) shows that $\hat{\gamma}(\mathbf{f})$ is a linear combination of these concave functions with positive coefficients, it follows that $\hat{\gamma}(\mathbf{f})$ is concave over Ω^+. This proves **Fact 2**. □

Sample Estimators of the Growth Rate Mean

Our proof of **Fact 3** uses the *Jensen inequality*. We will state and prove it before giving our proof.

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

$$g(\bar{z}) \leq \overline{g(z)} \qquad \left(\overline{g(z)} \leq g(\bar{z})\right), \tag{2.9}$$

where

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

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

Sample Estimators of the Growth Rate Mean

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 the inequality

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

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

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

The Jensen inequality then follows from the definitions of $\bar{z}$ and $\overline{g(z)}$. □

Sample Estimators of the Growth Rate Mean

Proof of Fact 3. Let $\mathbf{f} \in \Omega^+$. Then the points $\{r(d,\mathbf{f})\}_{d=1}^{D}$ all lie within an interval $[a,b] \subset (-1,\infty)$. Because $\log(1+r)$ is a concave function of r over $(-1,\infty)$, the Jensen inequality (2.9) and definition (2.8) of $\hat{\mu}(\mathbf{f})$ yield

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

This establishes the upper bound (2.7), whereby **Fact 3** is proved. □

Sample Estimators of the Growth Rate Mean

Remark. The Jensen inequality can yield other useful bounds. For example, if we take $g(z) = z^p$ for some $p > 1$, so that $g(z)$ is convex over $[0, \infty)$, and we take $z_d = w_d$ for every d then because the points $\{w_d\}_{d=1}^{D}$ all lie within $[0, 1]$, the Jensen inequality yields

$$\bar{w}^p = \left(\sum_{d=1}^{D} w_d^2 \right)^p \leq \sum_{d=1}^{D} w_d^{p+1} = \overline{w^p} .$$

Remark. Under very mild assumptions on the return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ that are always satisfied in practice we can strengthen **Fact 2** to

$$\hat{\gamma}(\mathbf{f}) \text{ is strictly concave over } \Omega^+ ,$$

and can strengthen bound (2.7) of **Fact 3** to the strict inequality

$$\hat{\gamma}(\mathbf{f}) < \log\big(1 + \hat{\mu}(\mathbf{f})\big) \quad \text{when } \mathbf{f} \neq \mathbf{0} . \tag{2.10}$$

Portfolios without Risk-Free Assets

Now we specialize to solvent Markowitz portfolios without risk-free assets. The associated allocations $\mathbf{f}$ belong to

$$\Omega = \left\{ \mathbf{f} \in \Omega^{+} \;:\; \mathbf{1}^{\mathrm{T}}\mathbf{f} = 1 \right\} . \tag{3.11}$$

On this set the growth rate mean sample estimator (2.6) reduces to

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

This is an infinitely differentiable function over Ω^{+} with

$$\begin{aligned} \nabla_{\mathbf{f}}\, \hat{\gamma}(\mathbf{f}) &= \sum_{d=1}^{D} w_d \, \frac{\mathbf{r}(d)}{1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}} , \\ \nabla_{\mathbf{f}}^{2} \hat{\gamma}(\mathbf{f}) &= - \sum_{d=1}^{D} w_d \, \frac{\mathbf{r}(d)\, \mathbf{r}(d)^{\mathrm{T}}}{(1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f})^{2}} . \end{aligned} \tag{3.13}$$

Portfolios without Risk-Free Assets

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

Fact 4. $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is nonpositive definite for every $\mathbf{f} \in \Omega$.

Fact 5. $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is negative definite for every $\mathbf{f} \in \Omega$ if and only if the vectors $\{\mathbf{r}(d)\}_{d=1}^{D}$ span $\mathbb{R}^N$.

Remark. Fact 4 implies that $\hat{\gamma}(\mathbf{f})$ is concave over Ω, which was already proven in **Fact 2**. **Fact 5** implies that $\hat{\gamma}(\mathbf{f})$ is strictly concave over Ω when the vectors $\{\mathbf{r}(d)\}_{d=1}^{D}$ span $\mathbb{R}^N$, which is always the case in practice.

Proof of Fact 4. Let $\mathbf{f} \in \Omega$. Then for every $\mathbf{y} \in \mathbb{R}^N$ we have

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

Therefore $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is nonpositive definite, which proves **Fact 4**. □

Portfolios without Risk-Free Assets

Proof of Fact 5. Let $\mathbf{f} \in \Omega$. Then by the calculation in the previous proof we see that for every $\mathbf{y} \in \mathbb{R}^N$

$$\mathbf{y}^{\mathrm{T}} \nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f}) \mathbf{y} = 0 \quad \iff \quad \mathbf{r}(d)^{\mathrm{T}} \mathbf{y} = 0 \ \ \forall d \, .$$

First, suppose that $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is not negative definite. Then there exists an $\mathbf{y} \in \mathbb{R}^N$ such that $\mathbf{y}^{\mathrm{T}} \nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f}) \mathbf{y} = 0$ and $\mathbf{y} \neq \mathbf{0}$. The vectors $\{\mathbf{r}(d)\}_{d=1}^D$ must then lie in the hyperplane orthogonal (normal) to $\mathbf{y}$. Therefore the vectors $\{\mathbf{r}(d)\}_{d=1}^D$ do not span $\mathbb{R}^N$.

Conversely, suppose that the vectors $\{\mathbf{r}(d)\}_{d=1}^D$ do not span $\mathbb{R}^N$. Then there must be a nonzero vector $\mathbf{y}$ that is orthogonal to their span. This means that $\mathbf{y}$ satisfies $\mathbf{r}(d)^{\mathrm{T}} \mathbf{y} = 0$ for every d, whereby $\mathbf{y}^{\mathrm{T}} \nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f}) \mathbf{y} = 0$. Therefore $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is not negative definite.

Both directions of the characterization in **Fact 5** are now proven. □

Portfolios without Risk-Free Assets

Henceforth we will assume that the covariance matrix $\mathbf{V}$ *is positive definite.* Recall that this is equivalent to assuming that the set $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ spans $\mathbb{R}^N$. Because this condition implies that the set $\{\mathbf{r}(d)\}_{d=1}^{D}$ spans $\mathbb{R}^N$, by **Fact 5** it implies that $\nabla_{\mathbf{f}}^2 \hat{\gamma}(\mathbf{f})$ is negative definite for every $\mathbf{f} \in \Omega$. *Therefore the estimator* $\hat{\gamma}(\mathbf{f})$ *is a strictly concave function over* Ω*.*

Remark. Because $\hat{\gamma}(\mathbf{f})$ is a strictly concave function over Ω, *if it has a maximum then it has a unique maximizer.* Indeed, suppose that $\hat{\gamma}(\mathbf{f})$ has maximum $\hat{\gamma}_{\mathrm{mx}}$ over Ω, and that $\mathbf{f}_0$ and $\mathbf{f}_1 \in \Omega$ are maximizers of $\hat{\gamma}(\mathbf{f})$ 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 $t \in (0,1)$ we have $\mathbf{f}_t \in \Omega$ and, by the strict concavity of $\hat{\gamma}(\mathbf{f})$ over Ω,

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

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

Portfolios without Risk-Free Assets

Recall that Ω^+ is the intersection of the half spaces

$$1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} > 0\,, \qquad \text{for } d = 1, \cdots, D\,,$$

and that Ω is the intersection of Ω^+ with the hyperplane $\mathbf{1}^{\mathrm{T}}\mathbf{f} = 1$.

For many return histories $\{\mathbf{r}(d)\}_{d=1}^{D}$ the set Ω is bounded. In that case we will have $1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} \searrow 0$ for at least one d as $\mathbf{f}$ approaches the boundary of Ω. But then we will have $\log(1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}) \to -\infty$ for at least one d as $\mathbf{f}$ approaches the boundary of Ω. Hence, when Ω is bounded we will have $\hat{\gamma}(\mathbf{f}) \to -\infty$ as $\mathbf{f}$ approaches the boundary of Ω. *Because $\hat{\gamma}(\mathbf{f})$ is continuous over Ω and goes to $-\infty$ as $\mathbf{f}$ approaches the boundary of Ω, when Ω is bounded $\hat{\gamma}(\mathbf{f})$ has a maximizer in Ω. Because $\hat{\gamma}(\mathbf{f})$ is strictly concave over Ω, this maximizer is unique.*

Mean-Variance Estimators of the Growth Rate Mean

The maximizer of $\hat{\gamma}(\mathbf{f})$ over Ω can be found numerically by methods that are typically covered in graduate courses. Rather than seek the maximizer of $\hat{\gamma}(\mathbf{f})$ over Ω, we will replace the estimator $\hat{\gamma}(\mathbf{f})$ with a new estimator for which finding the maximizer is easier. The hope is that the maximizer of $\hat{\gamma}(\mathbf{f})$ and the maximizer of the new estimator will be close.

This strategy rests upon the fact that $\hat{\gamma}(\mathbf{f})$ is itself an approximation. The uncertainties associated with it will translate into uncertainities about its maximizer. The hope is that the difference between the maximizer of $\hat{\gamma}(\mathbf{f})$ and that of the new estimator will be within these uncertainties.

Mean-Variance Estimators of the Growth Rate Mean

We will now derive some so-called *mean-variance estimators* for the growth rate mean $\gamma(\mathbf{f})$ of a Markowitz portfolio with allocation $\mathbf{f}$, each of which will have the form

$$\hat{\gamma}(\mathbf{f}) = G\big(\hat{\mu}(\mathbf{f}), \hat{\xi}(\mathbf{f})\big) ,$$

where $G(\mu, \xi)$ is some function, $\hat{\mu}(\mathbf{f})$ is the sample estimator of the return mean, and $\hat{\xi}(\mathbf{f})$ sample estimator of the return variance.

Remark. The fact that $\hat{\gamma}(\mathbf{f})$ is a function of $\hat{\mu}(\mathbf{f})$ and $\hat{\xi}(\mathbf{f})$ will mean that its maximizer can be found easily by using the efficient frontiers that we computed earlier.

For simplicity we will stay in the setting of solvent Markowitz portfolios without risk-free assets. In that case $\hat{\mu}(\mathbf{f})$ and $\hat{\xi}(\mathbf{f})$ are given by

$$\hat{\mu}(\mathbf{f}) = \mathbf{m}^{\mathrm{T}}\mathbf{f} , \qquad \hat{\xi}(\mathbf{f}) = \mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f} .$$

Mean-Variance Estimators of the Growth Rate Mean

A stratagy introduced by Markowitz in his 1959 book is to estimate $\hat{\gamma}(\mathbf{f})$ by using the *second-order Taylor approximation* of $\log(1+r)$ for small r. This approximation is

$$\log(1+r) \approx r - \tfrac{1}{2}r^2\,. \tag{4.14}$$

When this approximation is used in (3.12) we obtain the *quadratic estimator* of the growth rate mean

$$\begin{aligned}
\hat{\gamma}_{\mathrm{q}}(\mathbf{f}) &= \sum_{d=1}^{D} w_d \left(\mathbf{r}(d)^{\mathrm{T}}\mathbf{f} - \tfrac{1}{2}(\mathbf{r}(d)^{\mathrm{T}}\mathbf{f})^2\right) \\
&= \left(\sum_{d=1}^{D} w_d\, \mathbf{r}(d)\right)^{\mathrm{T}} \mathbf{f} - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}} \left(\sum_{d=1}^{D} w_d\, \mathbf{r}(d)\, \mathbf{r}(d)^{\mathrm{T}}\right) \mathbf{f} \\
&= \mathbf{m}^{\mathrm{T}}\mathbf{f} - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{m}\mathbf{m}^{\mathrm{T}} + \mathbf{V})\mathbf{f}\,.
\end{aligned} \tag{4.15}$$

Mean-Variance Estimators of the Growth Rate Mean

The second and third-order and Taylor approximation to $\log(1+r)$.

r	$\log(1+r)$	$r - \frac{1}{2}r^2$	$r - \frac{1}{2}r^2 + \frac{1}{3}r^3$
$-.5$	$-.69315$	$-.62500$	$-.66667$
$-.4$	$-.51083$	$-.48000$	$-.50133$
$-.3$	$-.35667$	$-.34500$	$-.35400$
$-.2$	$-.22314$	$-.22000$	$-.22267$
$-.1$	$-.10536$	$-.10500$	$-.10533$
$.0$	$.00000$	$.00000$	$.00000$
$.1$	$.09531$	$.09500$	$.09533$
$.2$	$.18232$	$.18000$	$.18267$
$.3$	$.26236$	$.25500$	$.26400$
$.4$	$.33647$	$.32000$	$.34133$
$.5$	$.40547$	$.37500$	$.41667$

Mean-Variance Estimators of the Growth Rate Mean

The table on the previous slide shows that the second-order Taylor approximation to $\log(1 + r)$ is pretty good when $|r| < .25$ and that it is not too bad when $.25 < |r| < .5$. It is bad when $|r| \geq .5$.

Remark. This observation suggests that the quadratic estimator $\hat{\gamma}_{\mathrm{q}}(\mathbf{f})$ given by (4.15) might only be trusted when the class of portfolio allocations being considered lies within

$$\Omega^{+}_{[.75,1.25]} = \left\{ \mathbf{f} \in \Omega^{+} \,:\, .75 \leq 1 + r(d, \mathbf{f}) \leq 1.25 \quad \forall d \right\}.$$

Such a restriction is usually satisfied by Λ, the set of long allocations, but is often not satisfied by highly leveraged portfolios. This is why good leveraged investors do not use the quadratic estimator.

Mean-Variance Estimators of the Growth Rate Mean

The *quadratic estimator* (4.15) can be expressed as

$$\hat{\gamma}_{\mathrm{q}}(\mathbf{f}) = \mathbf{m}^{\mathrm{T}}\mathbf{f} - \tfrac{1}{2}\left(\mathbf{m}^{\mathrm{T}}\mathbf{f}\right)^2 - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,. \tag{4.16}$$

We obtained this estimator twice earlier using the moment and cumulant generating functions.

Because it is often the case that

$$\left(\mathbf{m}^{\mathrm{T}}\mathbf{f}\right)^2 \quad \text{is much smaller than} \quad \mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,,$$

it is tempting to drop the $(\mathbf{m}^{\mathrm{T}}\mathbf{f})^2$ term in (4.16). This leads to the *parabolic estimator* of the growth rate mean

$$\hat{\gamma}_{\mathrm{p}}(\mathbf{f}) = \mathbf{m}^{\mathrm{T}}\mathbf{f} - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,. \tag{4.17}$$

Remark. While this estimator is commonly used, there are many times when it is not good. It is particularly bad in a bubble. We will see that using it can lead to overbetting at times when overbetting is very risky.

Mean-Variance Estimators of the Growth Rate Mean

We can also estimate $\hat{\gamma}(\mathbf{f})$ by the *mean-centered second-order Taylor* approximation of $\log(1+r)$ for $r = \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}$ near $\hat{\mu}(\mathbf{f}) = \mathbf{m}^{\mathrm{T}}\mathbf{f}$. That approximation is

$$\log(1+r) \approx \log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right) + \frac{(\mathbf{r}(d)-\mathbf{m})^{\mathrm{T}}\mathbf{f}}{1+\mathbf{m}^{\mathrm{T}}\mathbf{f}} - \frac{1}{2}\,\frac{\left((\mathbf{r}(d)-\mathbf{m})^{\mathrm{T}}\mathbf{f}\right)^2}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2}\,.$$

When this approximation is used in (3.12) we obtain the *Taylor estimator*

$$\hat{\gamma}_{\mathrm{t}}(\mathbf{f}) = \log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right) - \frac{1}{2}\,\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2}\,, \qquad (4.18)$$

which is defined over the half-space

$$H_{\mathrm{t}} = \left\{\mathbf{f} \in \mathbb{R}^N \,:\, 1+\mathbf{m}^{\mathrm{T}}\mathbf{f} > 0\right\}. \qquad (4.19)$$

This half-space contains Ω, the set of allocations for solvent Markowitz portfolios without risk-free assets.

Mean-Variance Estimators of the Growth Rate Mean

We obtained the Taylor estimator (4.18) earlier using the cumulant generating function. It satisfies the upper bound (2.7) from **Fact 3**. However, it is not concave over $H_{\rm t}$ and it does not generally have a maximum over $H_{\rm t}$. This makes it a poor replacement for $\hat{\gamma}(\mathbf{f})$ as an objective function over the entire set Ω^+, which is contained within $H_{\rm t}$.

The Hessian of $\hat{\gamma}_{\rm t}(\mathbf{f})$ over $H_{\rm t}$ is

$$\nabla_{\mathbf{f}}^2 \hat{\gamma}_{\rm t}(\mathbf{f}) = -\frac{\mathbf{m}\,\mathbf{m}^{\rm T} + \mathbf{V}}{(1+\mathbf{m}^{\rm T}\mathbf{f})^2} + 2\frac{\mathbf{V}\mathbf{f}\,\mathbf{m}^{\rm T} + \mathbf{m}\,\mathbf{f}^{\rm T}\mathbf{V}}{(1+\mathbf{m}^{\rm T}\mathbf{f})^3} - 3\frac{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}\,\mathbf{m}\,\mathbf{m}^{\rm T}}{(1+\mathbf{m}^{\rm T}\mathbf{f})^4}\,.$$

Mean-Variance Estimators of the Growth Rate Mean

This can be expressed as

$$
\begin{aligned}
\nabla_{\mathbf{f}}^2 \hat{\gamma}_t(\mathbf{f}) &= -\left(1 - \frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2}\right) \frac{\mathbf{m}\,\mathbf{m}^{\mathrm{T}}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2} \\
&\quad - \frac{\mathbf{V}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2} + 2\,\frac{\mathbf{V}\mathbf{f}\,\mathbf{m}^{\mathrm{T}} + \mathbf{m}\,\mathbf{f}^{\mathrm{T}}\mathbf{V}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^3} - 4\,\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,\mathbf{m}\,\mathbf{m}^{\mathrm{T}}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^4}\,. \\
&= -\left(1 - \frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2}\right) \frac{\mathbf{m}\,\mathbf{m}^{\mathrm{T}}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2} \\
&\quad - \left(\mathbf{I} - \frac{2\,\mathbf{f}\,\mathbf{m}^{\mathrm{T}}}{1+\mathbf{m}^{\mathrm{T}}\mathbf{f}}\right)^{\mathrm{T}} \frac{\mathbf{V}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2} \left(\mathbf{I} - \frac{2\,\mathbf{f}\,\mathbf{m}^{\mathrm{T}}}{1+\mathbf{m}^{\mathrm{T}}\mathbf{f}}\right).
\end{aligned}
$$

This is clearly negative definite over the set

$$
\Omega_t^+ = \left\{ \mathbf{f} \in \mathbb{R}^N \,:\, \sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}} \leq 1 + \mathbf{m}^{\mathrm{T}}\mathbf{f} \right\}.
$$

Mean-Variance Estimators of the Growth Rate Mean

Remark. The set Ω_{t}^{+} is convex. Indeed, if $\mathbf{f}_0$, $\mathbf{f}_1 \in \Omega_{\mathrm{t}}^{+}$ and we set $\mathbf{f}_t = (1-t)\mathbf{f}_0 + t\mathbf{f}_1$ then for every $t \in (0,1)$ we have

$$
\begin{aligned}
\mathbf{f}_t^{\mathrm{T}}\mathbf{V}\mathbf{f}_t &= (1-t)^2\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_0 + 2t(1-t)\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_1 + t^2\mathbf{f}_1^{\mathrm{T}}\mathbf{V}\mathbf{f}_1 \\
&\leq (1-t)^2\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_0 + 2t(1-t)\sqrt{\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_0}\sqrt{\mathbf{f}_1^{\mathrm{T}}\mathbf{V}\mathbf{f}_1} + t^2\mathbf{f}_1^{\mathrm{T}}\mathbf{V}\mathbf{f}_1 \\
&= \left((1-t)\sqrt{\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_0} + t\sqrt{\mathbf{f}_1^{\mathrm{T}}\mathbf{V}\mathbf{f}_1}\right)^2 \\
&\leq \Big((1-t)(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}_0) + t\,(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}_1)\Big)^2 = (1+\mathbf{m}^{\mathrm{T}}\mathbf{f}_t)^2 \,.
\end{aligned}
$$

The constraint for the set Ω_{t}^{+} can be expressed as a linear constraint $0 < 1 + \mathbf{m}^{\mathrm{T}}\mathbf{f}$ plus a quadratic constraint $\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f} \leq (1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2$. This quadratic constraint can become nondefinite and thereby harder to use.

Mean-Variance Estimators of the Growth Rate Mean

We now introduce an estimator with better properties that uses the first term from the Taylor estimator (4.18) and the volatility term from the quadratic estimator (4.16). This leads to the *reasonable estimator* of the growth rate mean

$$\hat{\gamma}_{\mathrm{r}}(\mathbf{f}) = \log\left(1 + \mathbf{m}^{\mathrm{T}}\mathbf{f}\right) - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,, \tag{4.20}$$

which is also defined over the half-space $H_{\mathrm{r}} = H_{\mathrm{t}}$ given by (4.19). This estimator is strictly concave and satisfies the upper bound (2.7) from **Fact 3** over H_{r}.

Mean-Variance Estimators of the Growth Rate Mean

Another growth rate mean estimator with good properties can be obtained by a different modification of (4.18) — namely, the sensible estimator

$$\hat{\gamma}_{\mathrm{s}}(\mathbf{f}) = \log\left(1 + \mathbf{m}^{\mathrm{T}}\mathbf{f}\right) - \frac{1}{2}\,\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{1 + 2\mathbf{m}^{\mathrm{T}}\mathbf{f}}\,, \tag{4.21}$$

which is defined on the half-space

$$H_{\mathrm{s}} = \left\{\mathbf{f} \in \mathbb{R}^N \,:\, 0 < 1 + 2\mathbf{m}^{\mathrm{T}}\mathbf{f}\right\}. \tag{4.22}$$

This half-space is smaller than the half-space H_{r} given by (4.19) over which the reasonable estimator (4.20) was defined.

The estimator $\hat{\gamma}_{\mathrm{s}}(\mathbf{f})$ clearly satisfies the upper bound (2.7) from **Fact 3** for every $\mathbf{f} \in H_{\mathrm{s}}$. Moreover, we have the following.

Fact 6. $\hat{\gamma}_{\mathrm{s}}(\mathbf{f})$ is strictly concave over the half-space H_{s}.

Mean-Variance Estimators of the Growth Rate Mean

Proof of Fact 6. We will show that $\hat{\gamma}_{\mathrm{s}}(\mathbf{f})$ is the sum of two functions, one of which is concave and the other of which is strictly concave over H_{s}.

The function $\log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right)$ is infinitely differentiable over H_{s} with

$$\nabla_{\mathbf{f}} \log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right) = \frac{\mathbf{m}}{1+\mathbf{m}^{\mathrm{T}}\mathbf{f}},$$

$$\nabla_{\mathbf{f}}^2 \log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right) = -\frac{\mathbf{m}\,\mathbf{m}^{\mathrm{T}}}{(1+\mathbf{m}^{\mathrm{T}}\mathbf{f})^2}.$$

Because its Hessian is nonpositive definite, the function $\log\left(1+\mathbf{m}^{\mathrm{T}}\mathbf{f}\right)$ is concave over H_{s}. The harder part of the proof of **Fact 6** is to show that

$$-\frac{1}{2}\,\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{1+2\mathbf{m}^{\mathrm{T}}\mathbf{f}} \quad \text{is strictly concave over } H_{\mathrm{s}}\,. \tag{4.23}$$

This will be shown using the next two facts, which we will state and prove before finishing the proof of **Fact 6**.

Mean-Variance Estimators of the Growth Rate Mean

Fact 7. Let $\mathbf{b}$, $\mathbf{x} \in \mathbb{R}^N$ such that $1 + \mathbf{b}^{\mathrm{T}}\mathbf{x} > 0$. Then $\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}$ is invertible with

$$\left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}\right)^{-1} = \mathbf{I} - \frac{\mathbf{x}\,\mathbf{b}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \,. \tag{4.24}$$

Proof of Fact 7. Just check that

$$\begin{aligned}
\left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}\right)\left(\mathbf{I} - \frac{\mathbf{x}\,\mathbf{b}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}}\right) &= \left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}\right) - \frac{\left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}\right)\mathbf{x}\,\mathbf{b}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \\
&= \mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}} - \frac{\mathbf{x}\,\mathbf{b}^{\mathrm{T}} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}}\mathbf{x}\,\mathbf{b}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \\
&= \mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}} - \frac{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}}\,\mathbf{x}\,\mathbf{b}^{\mathrm{T}} = \mathbf{I} \,.
\end{aligned}$$

The assertions of **Fact 7** then follow. □

Mean-Variance Estimators of the Growth Rate Mean

Fact 8. Let $\mathbf{A} \in \mathbb{R}^{N\times N}$ be symmetric and positive definite. Let $\mathbf{b} \in \mathbb{R}^N$. Let X be the half-space given by

$$X = \left\{\mathbf{x} \in \mathbb{R}^N \,:\, 1 + \mathbf{b}^{\mathrm{T}}\mathbf{x} > 0\right\}.$$

Then

$$\phi(\mathbf{x}) = \frac{1}{2}\,\frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \quad \text{is strictly convex over } X\,.$$

Proof of Fact 8. The function $\phi(\mathbf{x})$ is infinitely differentiable over X with

$$\nabla_{\mathbf{x}}\,\phi(\mathbf{x}) = \frac{\mathbf{A}\mathbf{x}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} - \frac{1}{2}\,\frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\,\mathbf{b}}{(1 + \mathbf{b}^{\mathrm{T}}\mathbf{x})^2}\,,$$

$$\nabla_{\mathbf{x}}^2\phi(\mathbf{x}) = \frac{\mathbf{A}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} - \frac{\mathbf{A}\mathbf{x}\,\mathbf{b}^{\mathrm{T}} + \mathbf{b}\,\mathbf{x}^{\mathrm{T}}\mathbf{A}}{(1 + \mathbf{b}^{\mathrm{T}}\mathbf{x})^2} + \frac{\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\,\mathbf{b}\,\mathbf{b}^{\mathrm{T}}}{(1 + \mathbf{b}^{\mathrm{T}}\mathbf{x})^3}\,.$$

Mean-Variance Estimators of the Growth Rate Mean

Then using (4.24) of **Fact 7** the Hessian can be expressed as

$$\nabla_{\mathbf{x}}^2 \phi(\mathbf{x}) = \left(\mathbf{I} - \frac{\mathbf{b}\,\mathbf{x}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \right) \frac{\mathbf{A}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \left(\mathbf{I} - \frac{\mathbf{x}\,\mathbf{b}^{\mathrm{T}}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \right)$$
$$= \left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}} \right)^{-\mathrm{T}} \frac{\mathbf{A}}{1 + \mathbf{b}^{\mathrm{T}}\mathbf{x}} \left(\mathbf{I} + \mathbf{x}\,\mathbf{b}^{\mathrm{T}} \right)^{-1} .$$

Because $\mathbf{A}$ is positive definite and $1 + \mathbf{b}^{\mathrm{T}}\mathbf{x} > 0$ for every $\mathbf{x} \in X$, this shows that $\nabla_{\mathbf{x}}^2 \phi(\mathbf{x})$ is positive definite for every $\mathbf{x} \in X$. Therefore $\phi(\mathbf{x})$ is strictly convex over X, thereby proving **Fact 8**. □

By setting $\mathbf{A} = \mathbf{V}$ and $\mathbf{b} = 2\mathbf{m}$ in **Fact 8** and using the fact that the negative of a strictly convex function is strictly concave, we establish (4.23), thereby completing the proof of **Fact 6**. □

Mean-Variance Estimators of the Growth Rate Mean

Finally, we identify a class of solvent Markowitz portfolios whose allocations lie within H_{s}.

Fact 9. $\Omega_{\frac{1}{2}} = \{\mathbf{f} \in \Omega \,:\, \frac{1}{2} \leq 1 + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f} \;\; \forall d\} \subset H_{\mathrm{s}}$.

Proof. Because $\Omega_{\frac{1}{2}} = \{\mathbf{f} \in \Omega \,:\, 0 \leq 1 + 2\mathbf{r}(d)^{\mathrm{T}}\mathbf{f} \;\; \forall d\}$, it is clear that $0 \leq 1 + 2\mathbf{m}^{\mathrm{T}}\mathbf{f}$ for every $\mathbf{f} \in \Omega_{\frac{1}{2}}$ with equality if only if $0 = 1 + 2\mathbf{r}(d)^{\mathrm{T}}\mathbf{f}$ for every d. But this implies that $(\mathbf{r}(d) - \mathbf{m})^{\mathrm{T}}\mathbf{f} = 0$ for every d, which implies that $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ does not span $\mathbb{R}^N$, which contradicts the assumption that $\mathbf{V}$ is positive definite. Hence, for every $\mathbf{f} \in \Omega_{\frac{1}{2}}$ we have $0 < 1 + 2\mathbf{m}^{\mathrm{T}}\mathbf{f}$, which implies that $\mathbf{f} \in H_{\mathrm{s}}$ by definition (4.22). Therefore $\Omega_{\frac{1}{2}} \subset H_{\mathrm{s}}$. □

Remark. This class excludes portfolios that would have dropped 50% in value during a single trading day over the history considered. This seems like a reasonable constraint for any long-term investor.

Portfolios with Risk-Free Assets

We now extend the estimators derived in the last section to solvent Markowitz portfolios with risk-free assets. Specifically, we will use the sample estimator $\hat{\gamma}(\mathbf{f})$ to derive new estimators of $\gamma(\mathbf{f})$ in terms of sample estimators of the return mean and variance given by

$$\hat{\mu}(\mathbf{f}) = \mu_{\text{rf}}(\mathbf{f})\big(1 - \mathbf{1}^{\text{T}}\mathbf{f}\big) + \mathbf{m}^{\text{T}}\mathbf{f}\,, \qquad \mathbf{f}^{\text{T}}\mathbf{V}\mathbf{f}\,, \tag{5.25}$$

where $\mathbf{m}$ and $\mathbf{V}$ are given by

$$\begin{aligned} \mathbf{m} &= \sum_{d=1}^{D} w_d \mathbf{r}(d)\,, \\ \mathbf{V} &= \sum_{d=1}^{D} w_d \big(\mathbf{r}(d) - \mathbf{m}\big)\big(\mathbf{r}(d) - \mathbf{m}\big)^{\text{T}}\,. \end{aligned} \tag{5.26}$$

These new return mean-variance estimators of $\gamma(\mathbf{f})$ will allow us to work within the framework of Markowitz portfolio theory.

Portfolios with Risk-Free Assets

We observe that $\hat{\mu}(\mathbf{f})$ is the sample mean of of the history $\{r(d,\mathbf{f})\}_{d=1}^{D}$ and that

$$r(d,\mathbf{f}) - \hat{\mu}(\mathbf{f}) = \tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}\,,$$

where $\tilde{\mathbf{r}}(d) = \mathbf{r}(d) - \mathbf{m}$. In words, $\tilde{\mathbf{r}}(d)$ is the deviation of $\mathbf{r}(d)$ from its sample mean $\mathbf{m}$. Then we can write

$$\begin{aligned}\log\big(1 + r(d,\mathbf{f})\big) &= \log\big(1 + \hat{\mu}(\mathbf{f})\big) + \frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})} \\ &\quad - \left(\frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})} - \log\left(1 + \frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})}\right)\right) .\end{aligned} \tag{5.27}$$

Notice that the last term on the first line has sample mean zero while the concavity of the function $r \mapsto \log(1 + r)$ implies that $r - \log(1 + r) \geq 0$, which implies that the term on the second line is nonpositive.

Portfolios with Risk-Free Assets

Therefore by taking the sample mean of (5.27) we obtain

$$\begin{aligned}\hat{\gamma}(\mathbf{f}) &= \sum_{d=1}^{D} w_d \log\big(1 + r(d,\mathbf{f})\big) \\ &= \log\big(1 + \hat{\mu}(\mathbf{f})\big) \\ &\quad - \sum_{d=1}^{D} w_d \left(\frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})} - \log\left(1 + \frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})}\right)\right) .\end{aligned} \tag{5.28}$$

The last sum will be positive whenever $\mathbf{f} \neq \mathbf{0}$ and $\mathbf{V}$ is positive definite.

Remark. By dropping the last term in the foregoing calculation we get an alternative proof of **Fact 3**, which was proved earlier using the Jensen inequality. Indeed, (5.28) can be viewed as an improvement upon **Fact 3**.

Portfolios with Risk-Free Assets

We can estimate $\hat{\gamma}(\mathbf{f})$ using the second-order Taylor approximation of $\log(1+r)$ for small r. This approximation is

$$\log(1+r) \approx r - \tfrac{1}{2}r^2 \,. \tag{5.29}$$

When this approximation is used inside the sum of (5.28) we obtain

$$\hat{\gamma}(\mathbf{f}) \approx \log\big(1+\hat{\mu}(\mathbf{f})\big) - \frac{1}{2}\sum_{d=1}^{D} w_d \left(\frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1+\hat{\mu}(\mathbf{f})}\right)^2 .$$

This leads to the mean-centered *Taylor estimator*

$$\hat{\gamma}_{\mathrm{t}}(\mathbf{f}) = \log\big(1+\hat{\mu}(\mathbf{f})\big) - \frac{1}{2}\,\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1+\hat{\mu}(\mathbf{f}))^2} \,. \tag{5.30}$$

While this estimator is defined over $1+\hat{\mu}(\mathbf{f}) > 0$, it is strictly concave over

$$\Omega_{\mathrm{t}}^{+} = \left\{ \mathbf{f} \in \mathbb{R}^N \,:\, \sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}} \leq 1+\hat{\mu}(\mathbf{f}) \right\} . \tag{5.31}$$

Portfolios with Risk-Free Assets

We can derive other estimators from the Taylor estimator.
The analog of the sensible estimator (4.21) is

$$\hat{\gamma}_{\rm s}(\mathbf{f}) = \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \frac{1}{2}\,\frac{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}{1 + 2\hat{\mu}(\mathbf{f})} \quad \text{over} \quad 1 + 2\hat{\mu}(\mathbf{f}) > 0\,. \tag{5.32}$$

The analog of the reasonable estimator (4.20) is

$$\hat{\gamma}_{\rm r}(\mathbf{f}) = \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \tfrac{1}{2}\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f} \qquad \text{over} \quad 1 + \hat{\mu}(\mathbf{f}) > 0\,. \tag{5.33}$$

The analog of the quadratic estimator (4.16) is

$$\hat{\gamma}_{\rm q}(\mathbf{f}) = \hat{\mu}(\mathbf{f}) - \tfrac{1}{2}\hat{\mu}(\mathbf{f})^2 - \tfrac{1}{2}\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f} \qquad \text{over} \quad \hat{\mu}(\mathbf{f}) \leq 1\,. \tag{5.34}$$

The analog of the parabolic estimator (4.17) is

$$\hat{\gamma}_{\rm p}(\mathbf{f}) = \hat{\mu}(\mathbf{f}) - \tfrac{1}{2}\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}\,. \tag{5.35}$$

Portfolios with Risk-Free Assets

The derivations of these estimators each assume that $|\hat{\mu}(\mathbf{f})| \ll 1$.

The sensible estimator (5.32) derives from the Taylor estimator (5.30) by dropping the $\hat{\mu}(\mathbf{f})^2$ term from the denominator $(1+\hat{\mu}(\mathbf{f}))^2 = 1 + 2\hat{\mu}(\mathbf{f}) + \hat{\mu}(\mathbf{f})^2$ under $\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}$.

The reasonable estimator (5.33) derives from the sensible estimator (5.32) by dropping the $2\hat{\mu}(\mathbf{f})$ term from the denominator under $\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}$.

The quadratic estimator (5.34) derives from the reasonable estimator (5.33) by replacing $\log(1+\hat{\mu}(\mathbf{f}))$ with the second-order Taylor approximation $\hat{\mu}(\mathbf{f}) - \frac{1}{2}\hat{\mu}(\mathbf{f})^2$. The result is an increasing function of $\hat{\mu}(\mathbf{f})$ when $\hat{\mu}(\mathbf{f}) \leq 1$.

The parabolic estimator (5.35) derives from the quadratic estimator (5.34) by also assuming that $\hat{\mu}(\mathbf{f})^2 \ll \mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}$ and dropping the $\hat{\mu}(\mathbf{f})^2$ term.

Portfolios with Risk-Free Assets

The Taylor, sensible, and reasonable estimators given by (5.30), (5.32), and (5.33) respectively each satisfy **Fact 1** and bound (2.7) from **Fact 3** over the set where it is defined.

The quadratic and parabolic estimators given by (5.34) and (5.35) respectively each satisfy analogs of **Fact 1** and bound (2.7) from **Fact 3** obtained by replacing the log by an appropriate Taylor approximation.

The sensible, reasonable, quadratic, and parabolic estimators given by (5.32), (5.33), (5.34), and (5.35) respectively are each strictly concave with a unique global maximum over the set where it is defined. These are analogs of **Fact 2**.

The Taylor estimator (5.30) is strictly concave with a unique maximum over the set Ω_t^+ given by (5.31). This is an analog of **Fact 2**.