Portfolios that Contain Risky Assets 15: Cautious Objectives for Markowitz Portfolios

C. David Levermore

University of Maryland, College Park, MD

Math 420: *Mathematical Modeling*
May 3, 2020 version
© 2020 Charles David Levermore

Portfolios that Contain Risky Assets
Part II: Stochastic Models

11. Independent, Identically-Distributed Models for Assets
12. Assessment of Independent, Identically-Distributed Models
13. Independent, Identically-Distributed Models for Portfolios
14. Kelly Objectives for Markowitz Portfolios
15. Cautious Objectives for Markowitz Portfolios
16. Optimization of Mean-Variance Objectives

Cautious Objectives for Markowitz Portfolios

1. Introduction
2. Cautious Objectives
3. Sample Mean Estimator Uncertainty
4. Central Limit Theorem
5. Downside Uncertainties
6. Mean-Variance Estimators of Cautious Objectives

Introduction

The Kelly criterion says that investors whose objective is to maximize the value of their portfolio over an extended period should maximize its growth rate mean. More precisely, it suggests that such investors should select an allocation $\mathbf{f}$ that maximizes the estimator $\hat{\gamma}(\mathbf{f})$. This suggestion rests upon:

- the validity of an IID model,
- the Law of Large Numbers,
- the accuracy of the estimator $\hat{\gamma}(\mathbf{f})$.

If these assumptions hold then the Kelly criterion would be suitable for many young investors, but not for those older investors who depend upon their portfolios for their income. Of course, the first assumption is questionable while the last is foolhardy, so even young investors should be more cautious.

Introduction

The Kelly criterion exposes investors to potential downside events from which it might be hard to recover. Older investors who depend upon their portfolios for their income might be drawing down on their portfolio at a rate of 4% per year, hoping that this income stream will last at least 20 years. But if the value of their portfolio is reduced by 40% in a market downturn then their future income stream will be similarly reduced. This puts them in a tough spot if their reduced income no longer covers their fixed expenses. They might feel forced to draw down at 6.5% per year, which would rapidly erode the value of their portfolio.

Here we will develop objective functions that are better suited for more cautious investors. We will do so within the framework of IID models.

Cautious Objectives

Given a set of assets with a return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ and a choice of positive weights $\{w_d\}_{d=1}^{D}$ that sum to one, the Kelly Criterion selects the portfolio allocation $\mathbf{f}$ that maximizes $\hat{\gamma}(\mathbf{f})$ over a class Π of Markowitz portfolio allocations $\mathbf{f}$, where the objective $\hat{\gamma}(\mathbf{f})$ is the sample mean estimator of $\gamma(\mathbf{f})$ given by

$$\hat{\gamma}(\mathbf{f}) = \sum_{d=1}^{D} w_d \log\big(1 + r(d,\mathbf{f})\big)\,, \tag{2.1}$$

with

$$r(d,\mathbf{f}) = \mu_{\mathrm{rf}}(\mathbf{f})(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}) + \mathbf{r}(d)^{\mathrm{T}}\mathbf{f}\,,$$
$$\mu_{\mathrm{rf}}(\mathbf{f}) = \begin{cases} \mu_{\mathrm{si}} & \text{for } \mathbf{1}^{\mathrm{T}}\mathbf{f} \le 1\,, \\ \mu_{\mathrm{cl}} & \text{for } \mathbf{1}^{\mathrm{T}}\mathbf{f} > 1\,. \end{cases} \tag{2.2}$$

Cautious Objectives

Here we present the family of *cautious objectives* that has the form

$$\widehat{\Gamma}^{\chi}(\mathbf{f}) = \hat{\gamma}(\mathbf{f}) - \chi \sqrt{\hat{\theta}(\mathbf{f})}\,, \tag{2.3}$$

where the nonnegative parameter χ is the so-called *caution coefficient* and $\hat{\theta}(\mathbf{f})$ is the sample variance estimator of $\theta(\mathbf{f})$ given by

$$\hat{\theta}(\mathbf{f}) = \frac{1}{1 - \bar{w}_D} \sum_{d=1}^{D} w_d \Big(\log\big(1 + r(d, \mathbf{f})\big) - \hat{\gamma}(\mathbf{f}) \Big)^2 \,, \tag{2.4}$$

with $\bar{w}_D$ defined by

$$\bar{w}_D = \sum_{d=1}^{D} w_d^2 \,. \tag{2.5}$$

Cautious Objectives

Because $\hat{\gamma}(\mathbf{f})$ is a strictly concave function of $\mathbf{f}$, it is clear that $\widehat{\Gamma}^{\chi}(\mathbf{f})$ given by (2.3) will be is a strictly concave function of $\mathbf{f}$ over any bounded set provided that the caution coefficient χ is small enough.

We will soon see that the function $\sqrt{\hat{\theta}(\mathbf{f})}$ is convex over relevant sets of $\mathbf{f}$. In that case for every $\chi > 0$ the additional term in $\widehat{\Gamma}^{\chi}(\mathbf{f})$ will enhance the strict concavity of $\hat{\gamma}(\mathbf{f})$, which helps guard against overbetting.

The choice of a value for the caution coefficient χ is up to each individual investor. It characterizes how cautious the investor wishes to be. Caution can arise from many sources, each of which has to be quantified in order to guide the choice of χ.

Cautious Objectives

We will consider contributions to the caution coefficient from two sources:

- the uncertainty in the sample mean estimator $\hat{\gamma}(\mathbf{f})$ given by (2.1);
- the desire to reduce the impact of downside market events.

The first will be analyzed using the Chebyshev inequality bounds that we developed earlier.

The second will require more information than the Law of Large Numbers provides. However, this additional information can be estimated with the aid of the *Central Limit Theorem.*

Other potential contributions to the caution coefficient will be explored in the projects. These might include our confidence in the IID model or our assesment of economic factors.

Sample Mean Estimator Uncertainty

Recall that the uncertainty in the sample mean estimator $\hat{\gamma}(\mathbf{f})$ can be quantified by the Chebyshev inequality, which shows for every $\delta > \sqrt{\bar{w}_D}$ that

$$\mathrm{Pr}\left\{ |\hat{\gamma}(\mathbf{f}) - \gamma(\mathbf{f})| > \delta \sqrt{\theta(\mathbf{f})} \right\} \leq \frac{\bar{w}_D}{\delta^2} .$$

This can be recast as

$$\mathrm{Pr}\left\{ |\hat{\gamma}(\mathbf{f}) - \gamma(\mathbf{f})| \leq \delta \sqrt{\theta(\mathbf{f})} \right\} \geq 1 - \frac{\bar{w}_D}{\delta^2} . \tag{3.6}$$

This implies that

$$\mathrm{Pr}\left\{ \gamma(\mathbf{f}) \geq \hat{\gamma}(\mathbf{f}) - \delta \sqrt{\theta(\mathbf{f})} \right\} \geq 1 - \frac{\bar{w}_D}{\delta^2} . \tag{3.7}$$

Sample Mean Estimator Uncertainty

This suggests that if $\bar{w}_D/\delta^2$ is small then with high probability

$$\gamma(\mathbf{f}) \geq \hat{\gamma}(\mathbf{f}) - \delta \sqrt{\hat{\theta}(\mathbf{f})}\,. \tag{3.8}$$

Notice that here we have replaced the unknown $\theta(\mathbf{f})$ in (3.7) with its estimator $\hat{\theta}(\mathbf{f})$ given by (2.4).

Remark. Because we see from (2.2) that $r(d, \mathbf{0}) = \mu_{\mathrm{si}}$ for every d, we see from (2.1) and (2.4) that

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

Therefore because $\gamma(\mathbf{0}) = \log(1 + \mu_{\mathrm{si}})$, we see that inequality (3.8) is an equality for $\mathbf{f} = \mathbf{0}$.

Sample Mean Estimator Uncertainty

Now let $\lambda_{\mathrm{e}} \in (0,1)$ be the probability that we hope inequality (3.8) holds. By setting

$$\lambda_{\mathrm{e}} = 1 - \frac{\bar{w}_D}{\delta^2} \,,$$

we obtain

$$\delta = \sqrt{\frac{\bar{w}_D}{1 - \lambda_{\mathrm{e}}}} \,.$$

This suggests that if this sample mean estimator uncertainty was our only concern then we could select the caution coefficient

$$\chi_{\mathrm{e}} = \sqrt{\frac{\bar{w}_D}{1 - \lambda_{\mathrm{e}}}} \,. \qquad (3.9)$$

Of course, the addition of other concerns will lead to a higher value for χ.

Central Limit Theorem

Because the Central Limit Theorem is used to analyze our next concern, we now review it. Let $\{X_d\}_{d=1}^{\infty}$ be any sequence of IID random variables drawn from a probability density $p(X)$ with mean γ and variance $\theta > 0$. Let $\{Y_d\}_{d=1}^{\infty}$ be the sequence of random variables defined by

$$Y_d = \frac{1}{d} \sum_{d'=1}^{d} X_{d'} \qquad \text{for every } d = 1, \cdots, \infty\,. \tag{4.10}$$

Recall that

$$\mathrm{Ex}(Y_d) = \gamma\,, \qquad \mathrm{Var}(Y_d) = \frac{\theta}{d}\,.$$

The Law of Large Numbers says that Y_d is will approach γ as $d \to \infty$. However, it does not say how the Y_d are distibuted around γ for fixed d. The Central Limit Theorem gives such information.

Central Limit Theorem

Let $\{Z_d\}_{d=1}^{\infty}$ be the sequence of random variables defined by

$$Z_d = \frac{Y_d - \gamma}{\sqrt{\theta/d}} \qquad \text{for every } d = 1, \cdots, \infty\,.$$

These random variables have been standardized so that

$$\mathrm{Ex}(Z_d) = 0\,, \qquad \mathrm{Var}(Z_d) = 1\,.$$

The Central Limit Theorem says that as $d \to \infty$ the limiting distribution of Z_d will be the mean-zero, variance-one normal distribution.

Central Limit Theorem

More precisely, it says that for every $\zeta \in \mathbb{R}$ we have

$$\lim_{d \to \infty} \mathrm{Pr}\{Z_d \geq -\zeta\} = \mathrm{N}(\zeta) , \tag{4.11}$$

where $\mathrm{N}(\zeta)$ is the normal cummulative distribution function defined by

$$\mathrm{N}(\zeta) \equiv \int_{-\infty}^{\zeta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \, \mathrm{d}Z = \int_{-\zeta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2} \, \mathrm{d}Z . \tag{4.12}$$

We can express the limit (4.11) in terms of Y_d as

$$\lim_{d \to \infty} \mathrm{Pr}\left\{ Y_d \geq \gamma - \zeta\sqrt{\theta/d} \right\} = \mathrm{N}(\zeta) . \tag{4.13}$$

Central Limit Theorem

Remark. The normal cummulative distribution function N is an increasing, continous function that maps $\mathbb{R}$ onto $(0,1)$. It thereby has an inverse N^{-1} that is an increasing, continuous function that maps $(0,1)$ onto $\mathbb{R}$. Both of these functions are infinitely differentiable.

Remark. The power of the Central Limit Theorem is that it assumes so little about the underlying probability density $p(X)$. Specifically, it assumes that

$$\int_{-\infty}^{\infty} X^2 p(X)\,\mathrm{d}X < \infty\,,$$

and that

$$0 < \theta = \int_{-\infty}^{\infty} (X-\gamma)^2 p(X)\,\mathrm{d}X\,, \quad \text{where} \quad \gamma = \int_{-\infty}^{\infty} X\,p(X)\,\mathrm{d}X\,.$$

Central Limit Theorem

The Central Limit Theorem does not estimate how fast the limit (4.13) is approached. Such estimates require additional assumptions about the underlying probability density $p(X)$. The Berry-Esseen Theorem is the simplest such theorem, but it is not covered in most undergraduate probability courses.

The Berry-Esseen Theorem says that there exists $C_{\mathrm{BE}} \in \mathbb{R}$ such that if

$$\rho = \int_{-\infty}^{\infty} |X - \gamma|^3 p(X)\,\mathrm{d}X < \infty\,,$$

then for every $\zeta \in \mathbb{R}$ we have

$$\left|\mathrm{Pr}\left\{ Y_d \geq \gamma - \zeta\sqrt{\theta/d} \right\} - \mathrm{N}(\zeta)\right| \leq C_{\mathrm{BE}}\,\frac{\rho}{\sqrt{d\,\theta^3}}\,. \tag{4.14}$$

Central Limit Theorem

Remark. The Berry-Esseen Theorem shows that the rate of convergence of the limit (4.13) in the Central Limit Theorem is $d^{-\frac{1}{2}}$ as $d \to \infty$.

Remark. The constant C_{BE} is universal because it does not depend upon the probability density $p(X)$. Its best value is known to lie within the interval $(0.4, 0.5)$. Bounding this value is the subject of current research.

Remark. The quantity ρ is the absolute centered third moment of the probability density $p(X)$. The Hölder inequality shows that it is bounded below by the variance θ of $p(X)$ as

$$\rho \geq \sqrt{\theta^3}\,.$$

The error bound (4.14) depends upon $p(X)$ through the ratio $\rho/\sqrt{\theta^3}$, which is larger for densities with fatter tails.

Downside Uncertainties

The IID model for the Markowitz portfolio with allocation $\mathbf{f}$ has a growth rate mean $\gamma(\mathbf{f})$ and a growth rate variance $\theta(\mathbf{f})$ that are estimated from the return history $\{\mathbf{r}(d)\}_{d=1}^{D}$ by $\hat{\gamma}(\mathbf{f})$ and $\hat{\theta}(\mathbf{f})$ given by (2.1) and (2.4).

Let $\{X_d\}_{d=1}^{\infty}$ be an IID growth rate history drawn from this model and let $\{Y_d\}_{d=1}^{\infty}$ be defined by (4.10). The Law of Large Numbers says that as $d \to \infty$ the values of Y_d become strongly peak around $\gamma(\mathbf{f})$. This behavior seems to be consistent with the idea that a reasonable approach towards portfolio management is to select $\mathbf{f}$ to maximize the estimator $\hat{\gamma}(\mathbf{f})$. However, by taking $\zeta = 0$ in (4.13) we see that the Central Limit Theorem implies

$$\lim_{d\to\infty} \mathrm{Pr}\{ Y_d \geq \gamma(\mathbf{f})\} = \mathrm{N}(0) = \tfrac{1}{2}\,.$$

This shows that in the long run the growth rate of a portfolio will exceed $\gamma(\mathbf{f})$ with a probability of only $\frac{1}{2}$. Cautious investors might want the portfolio to exceed the optimized growth rate with a higher probability.

Downside Uncertainties

The Central Limit Theorem says that if T is large enough then we can use the approximation

$$\mathrm{Pr}\left\{ Y_T \geq \gamma(\mathbf{f}) - \zeta\sqrt{\theta(\mathbf{f})/T} \right\} \approx \mathrm{N}(\zeta)\,. \tag{5.15}$$

Let $\lambda_\mathrm{d} \in (\frac{1}{2}, 1)$ be the probability that we do not want to experience a downside event. Set

$$\zeta_\mathrm{d} = \mathrm{N}^{-1}(\lambda_\mathrm{d})\,.$$

Then approximation (5.15) becomes

$$\mathrm{Pr}\left\{ Y_T \geq \gamma(\mathbf{f}) - \frac{\zeta_\mathrm{d}}{\sqrt{T}}\sqrt{\theta} \right\} \approx \lambda_\mathrm{d}\,. \tag{5.16}$$

Downside Uncertainties

This suggests that if downside tail events were our only concern then we could pick the caution coefficient

$$\chi_{\rm d} = \frac{{\rm N}^{-1}(\lambda_{\rm d})}{\sqrt{T}}\,, \tag{5.17}$$

whereby

$$\widehat{\Gamma}^{\chi}(\mathbf{f}) = \hat{\gamma}(\mathbf{f}) - \frac{{\rm N}^{-1}(\lambda_{\rm d})}{\sqrt{T}}\,\sqrt{\hat{\theta}(\mathbf{f})}\,. \tag{5.18}$$

Remark. Because $\hat{\gamma}(\mathbf{f})$ is a strictly concave function of $\mathbf{f}$, it is clear that $\widehat{\Gamma}^{\chi}(\mathbf{f})$ will be is a strictly concave function of $\mathbf{f}$ over any bounded set provided that $\chi_{\rm d}$ is small enough.

Downside Uncertainties

Remark. *The only assumption that we made made beyond the validity of the IID model in order to construct this objective is that T is large enough for the Central Limit Theorem to yield a good approximation of the distribution of growth rates.*

Remark. Investors often choose T to be the interval at which the portfolio will be rebalanced, regardless of whether T is large enough for the Central Limit Theorem approximation to be valid. If an investor plans to rebalance once a year then $T = 252$, twice a year then $T = 126$, four times a year then $T = 63$, and twelve times a year then $T = 21$. The smaller T, the less likely it is that the Central Limit Theorem approximation is valid.

Downside Uncertainties

The idea will be to select the admissible Markowitz allocation $\mathbf{f}$ *that maximizes* $\widehat{\Gamma}^{\chi}(\mathbf{f})$ *given a choice of* χ_{d} *by the investor. In other words, the objective will be to maximize the growth rate that will be exceeded by the portfolio with probability* λ_{d} *when it is held for* T *trading days.* Because $1 - \lambda_{\mathrm{d}}$ is the fraction of times the investor is willing to experience a downside tail event, the choice of λ_{d} reflects the caution of the investor. More cautious investors will select a higher λ_{d}.

Remark. The caution of an investor can increase with age. Retirees whose portfolio provide an income that covers much of their living expenses will often be extremely cautious. Investors within ten years of retirement may be fairly cautious because they have less time for their portfolio to recover from an economic downturn. In constrast, young investors can be less cautious because they have more time to experience economic upturns and because they are typically far from their peak earning capacity.

Downside Uncertainties

Remark. The caution of an investor should also depend on a careful reading of economic factors or an analysis of the historical data. For example, if the historical data shows evidence of a bubble then any investor should be more cautious.

An investor can select $\zeta_{\mathrm{d}} = \mathrm{N}^{-1}(\lambda_{\mathrm{d}})$ such that λ_{d} is a probability that reflects his or her caution. For example, an investor can select ζ_{d} based on the following tabulations

$$\begin{aligned}
&\mathrm{N}\left(\tfrac{1}{4}\right) \approx .5987\,, \quad \mathrm{N}\left(\tfrac{1}{2}\right) \approx .6915\,, \quad \mathrm{N}\left(\tfrac{3}{4}\right) \approx .7734\,, \quad \mathrm{N}(1) \approx .8413\,,\\
&\mathrm{N}\left(\tfrac{5}{4}\right) \approx .8944\,, \quad \mathrm{N}\left(\tfrac{3}{2}\right) \approx .9332\,, \quad \mathrm{N}\left(\tfrac{7}{4}\right) \approx .9599\,, \quad \mathrm{N}(2) \approx .9772\,,\\
&\mathrm{N}\left(\tfrac{9}{4}\right) \approx .9878\,, \quad \mathrm{N}\left(\tfrac{5}{2}\right) \approx .9938\,, \quad \mathrm{N}\left(\tfrac{11}{4}\right) \approx .9970\,, \quad \mathrm{N}(3) \approx .9987\,.
\end{aligned}$$

Downside Uncertainties

An investor who is willing to experience a downside tail event roughly

$$
\begin{aligned}
\text{once every two years might select } \zeta_{\rm d} &= 0\,, \\
\text{twice every five years might select } \zeta_{\rm d} &= \tfrac{1}{4}\,, \\
\text{thrice every ten years might select } \zeta_{\rm d} &= \tfrac{1}{2}\,, \\
\text{twice every nine years might select } \zeta_{\rm d} &= \tfrac{3}{4}\,, \\
\text{once every six years might select } \zeta_{\rm d} &= 1\,, \\
\text{once every ten years might select } \zeta_{\rm d} &= \tfrac{5}{4}\,, \\
\text{once every fifteen years might select } \zeta_{\rm d} &= \tfrac{3}{2}\,, \\
\text{once every twenty five years might select } \zeta_{\rm d} &= \tfrac{7}{4}\,, \\
\text{once every forty four years might select } \zeta_{\rm d} &= 2\,.
\end{aligned}
$$

Downside Uncertainties

Remark. *We should pick a larger value of ζ_{d} whenever our analysis of the historical data gives us less confidence either in the health of the economy, in the calibration of* $\mathbf{m}$ *and* $\mathbf{V}$*, or in the validity of an IID model.* These are the questions that are addressed in the projects.

Remark. This approach is similar to something in financal management called *value at risk*. The finance problem is much harder because the time horizon T considered there is much shorter, typically on the order of days. The Central Limit Theorem approximation is likely invalid in that case.

Mean-Variance Estimators of Cautious Objectives

The family of cautious objectives $\widehat{\Gamma}^{\chi}(\mathbf{f})$ given by (2.3) is expressed in terms of the growth rate mean and variance estimators $\hat{\gamma}(\mathbf{f})$ and $\hat{\theta}(\mathbf{f})$. In order to work within the framework of Markowitz portfolio theory we now derive estimators of these objectives expressed in terms of the sample estimators of the return mean and variance given by

$$\hat{\mu}(\mathbf{f}) = \mu_{\mathrm{rf}}(\mathbf{f})(1 - \mathbf{1}^{\mathrm{T}}\mathbf{f}) + \mathbf{m}^{\mathrm{T}}\mathbf{f}\,, \qquad \mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,, \tag{6.19}$$

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

$$\begin{aligned} \mathbf{m} &= \sum_{d=1}^{D} w_d\, \mathbf{r}(d)\,, \\ \mathbf{V} &= \sum_{d=1}^{D} w_d\, \big(\mathbf{r}(d) - \mathbf{m}\big)\big(\mathbf{r}(d) - \mathbf{m}\big)^{\mathrm{T}}\,. \end{aligned} \tag{6.20}$$

These are mean-variance estimators of the cautious objectives.

Mean-Variance Estimators of Cautious Objectives

Let $\tilde{\mathbf{r}}(d) = \mathbf{r}(d) - \mathbf{m}$ be the deviation of $\mathbf{r}(d)$ from its sample mean $\mathbf{m}$. Then

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

where $\hat{\mu}(\mathbf{f})$ is the sample mean given by (6.19) of the history $\{r(d,\mathbf{f})\}_{d=1}^{D}$. Then we can write

$$\begin{aligned} x(d,\mathbf{f}) &= \log\big(1 + r(d,\mathbf{f})\big) \\ &= \log\big(1 + \hat{\mu}(\mathbf{f})\big) + \log\left(1 + \frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})}\right) . \end{aligned} \tag{6.21}$$

In the last lecture this expression was used to derive estimators of $\hat{\gamma}(\mathbf{f})$. Here it will be used to also derive estimators of $\hat{\theta}(\mathbf{f})$.

Mean-Variance Estimators of Cautious Objectives

More specifically, in the last lecture we used the second-order Taylor approximation $\log(1+z) \approx z - \frac{1}{2}z^2$ in the second term of (6.21) to obtain

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

This led to the Taylor estimator

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

This estimator is not well-behaved for large $\mathbf{f}$, so from it we derived the sensible, reasonable, quadratic, and parabolic estimators, $\hat{\gamma}_{\mathrm{s}}(\mathbf{f})$, $\hat{\gamma}_{\mathrm{r}}(\mathbf{f})$, $\hat{\gamma}_{\mathrm{q}}(\mathbf{f})$, and $\hat{\gamma}_{\mathrm{p}}(\mathbf{f})$, all of which behave better for large $\mathbf{f}$.

Mean-Variance Estimators of Cautious Objectives

Here we use the first-order Taylor approximation $\log(1+z) \approx z$ in the second term of (6.21) to obtain

$$x(d, \mathbf{f}) \approx \log\big(1 + \hat{\mu}(\mathbf{f})\big) + \frac{\tilde{\mathbf{r}}(d)^{\mathrm{T}}\mathbf{f}}{1 + \hat{\mu}(\mathbf{f})} . \tag{6.22}$$

When this is placed into the definition of $\hat{\theta}(\mathbf{f})$ we obtain

$$\hat{\theta}(\mathbf{f}) = \sum_{d=1}^{D} \frac{w_d}{1 - \bar{w}} \big(x(d, \mathbf{f}) - \hat{\gamma}(\mathbf{f})\big)^2 \approx \frac{1}{1 - \bar{w}} \frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1 + \hat{\mu}(\mathbf{f}))^2} ,$$

which leads to the Taylor estimator

$$\hat{\theta}_{\mathrm{t}}(\mathbf{f}) = \frac{1}{1 - \bar{w}} \frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1 + \hat{\mu}(\mathbf{f}))^2} . \tag{6.23}$$

Like the Taylor estimator $\hat{\gamma}_{\mathrm{t}}(\mathbf{f})$, this is not well-behaved for large $\mathbf{f}$.

Mean-Variance Estimators of Cautious Objectives

The simplest thing to do is drop the $\hat{\mu}(\mathbf{f})$ term in the denominator of $\hat{\theta}_{\mathrm{t}}(\mathbf{f})$, which leads to the quadratic estimator

$$\hat{\theta}_{\mathrm{q}}(\mathbf{f}) = \frac{1}{1-\bar{w}}\,\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}\,. \tag{6.24}$$

We then introduce the *caution coefficient* χ by

$$\begin{aligned}\chi &= \frac{1}{\sqrt{1-\bar{w}}}\left(\chi_{\mathrm{e}}+\chi_{\mathrm{d}}\right)\\ &= \frac{1}{\sqrt{1-\bar{w}}}\left(\sqrt{\frac{\bar{w}}{1-\lambda_{\mathrm{e}}}}+\frac{\mathrm{N}^{-1}(\lambda_{\mathrm{d}})}{\sqrt{T}}\right)\,.\end{aligned} \tag{6.25}$$

Typically $\chi < 1$.

Mean-Variance Estimators of Cautious Objectives

When the estimator (6.24) is combined with the parabolic estimator $\hat{\gamma}_{\mathrm{p}}(\mathbf{f})$ we obtain

$$\widehat{\Gamma}_{\mathrm{p}}^{\chi}(\mathbf{f}) = \hat{\mu}(\mathbf{f}) - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f} - \chi\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}\,. \tag{6.26}$$

When it is combined with the quadratic estimator $\hat{\gamma}_{\mathrm{q}}(\mathbf{f})$ we obtain

$$\begin{aligned} \widehat{\Gamma}_{\mathrm{q}}^{\chi}(\mathbf{f}) &= \hat{\mu}(\mathbf{f}) - \tfrac{1}{2}\hat{\mu}(\mathbf{f})^2 - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f} - \chi\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}\,, \\ \text{over} \quad \Pi_{\mathrm{q}} &= \left\{\mathbf{f} \in \mathbb{R}^N \,:\, \hat{\mu}(\mathbf{f}) \leq 1\right\}. \end{aligned} \tag{6.27}$$

When it is combined with the reasonable estimator $\hat{\gamma}_{\mathrm{r}}(\mathbf{f})$ we obtain

$$\begin{aligned} \widehat{\Gamma}_{\mathrm{r}}^{\chi}(\mathbf{f}) &= \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \tfrac{1}{2}\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f} - \chi\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}\,, \\ \text{over} \quad \Pi_{\mathrm{r}} &= \left\{\mathbf{f} \in \mathbb{R}^N \,:\, 1 + \hat{\mu}(\mathbf{f}) > 0\right\}. \end{aligned} \tag{6.28}$$

Mean-Variance Estimators of Cautious Objectives

When it is combined with the sensible estimator $\hat{\gamma}_{\rm s}(\mathbf{f})$ we obtain

$$\widehat{\Gamma}^{\chi}_{\rm s}(\mathbf{f}) = \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \tfrac{1}{2}\frac{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}{1 + 2\hat{\mu}(\mathbf{f})} - \chi\sqrt{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}\,, \quad \text{over} \quad \Pi_{\rm s} = \left\{\mathbf{f} \in \mathbb{R}^N \,:\, 1 + \hat{\mu}(\mathbf{f}) > \tfrac{1}{2}\right\}. \tag{6.29}$$

When it is combined with the Taylor estimator $\hat{\gamma}_{\rm t}(\mathbf{f})$ we obtain

$$\widehat{\Gamma}^{\chi}_{\rm t}(\mathbf{f}) = \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \tfrac{1}{2}\frac{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}{(1 + \hat{\mu}(\mathbf{f}))^2} - \chi\sqrt{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}\,, \quad \text{over} \quad \Pi_{\rm t} = \left\{\mathbf{f} \in \mathbb{R}^N \,:\, 1 + \hat{\mu}(\mathbf{f}) \geq \sqrt{\mathbf{f}^{\rm T}\mathbf{V}\mathbf{f}}\right\}. \tag{6.30}$$

This estimator is strictly concave over $\Pi_{\rm t}$.

Mean-Variance Estimators of Cautious Objectives

Finally, when $\chi < 1$ and the Taylor estimator $\hat{\gamma}_{\mathrm{t}}(\mathbf{f})$ is combined with the Taylor estimator $\hat{\theta}_{\mathrm{t}}(\mathbf{f})$ given by (6.23) to estimate the cautious objective, $\widehat{\Gamma}^{\chi}(\mathbf{f})$ given by (2.3), then we obtain the ultimate estimator

$$\begin{aligned} \widehat{\Gamma}^{\chi}_{\mathrm{u}}(\mathbf{f}) &= \log\big(1 + \hat{\mu}(\mathbf{f})\big) - \tfrac{1}{2}\frac{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}{(1 + \hat{\mu}(\mathbf{f}))^2} - \chi\,\frac{\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}}{1 + \hat{\mu}(\mathbf{f})}\,, \\ \text{over} \quad \Pi^{\chi}_{\mathrm{u}} &= \left\{ \mathbf{f} \in \mathbb{R}^N \,:\, 1 + \hat{\mu}(\mathbf{f}) \geq \frac{\sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\mathbf{f}}}{1 - \chi} \right\}. \end{aligned} \tag{6.31}$$

This estimator is strictly concave over Π^{χ}_{u}.