Portfolios that Contain Risky Assets
10.3. Optimization of Mean-Variance Objectives

C. David Levermore

University of Maryland, College Park, MD

Math 420: *Mathematical Modeling*
April 24, 2022 version

Portfolios that Contain Risky Assets
Part II: Probabilistic Models

Portfolios that Contain Risky Assets
Part II: Probabilistic Models

10. Cautious Objectives for Portfolio Models

Optimization of Mean-Variance Objectives

Mean-Variance Objectives (Introduction)

Here we address the maximization problem for a mean-variance objective $\widehat{\Gamma}$ defined over a convex set Π of Markowitz allocations. These objectives have the general form

$$\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu}) \,, \tag{1.1a}$$

where

- $\hat{\sigma}$ is the volatility estimator defined over Π,
- $\hat{\mu}$ is the return mean estimator defined over Π,

and $G(\sigma, \mu)$ is defined over a set Σ_G of the $\sigma\mu$-plane that satisfies

$$\Sigma_G \supset \Sigma(\Pi) = \Big\{ (\hat{\sigma}, \hat{\mu}) \,:\, \text{all allocations in } \Pi \Big\} . \tag{1.1b}$$

Additional requirements will be imposed upon both $G(\sigma, \mu)$ and Σ_G in order to solve the the maximization problem.

Mean-Variance Objectives (Examples in $\mathcal{M}_+$)

Recall that:

- $\hat{\sigma}$ is convex function of the allocations;
- $\hat{\mu}$ is an affine function of the allocations.

We illustrate this with examples.

When $\Pi \subset \mathcal{M}_+$ we have

$$\hat{\sigma}(\mathbf{f}) = \sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\,\mathbf{f}}\,, \qquad \hat{\mu}(\mathbf{f}) = \mu_{\mathrm{rf}} + (\mathbf{m} - \mu_{\mathrm{rf}}\mathbf{1})^{\mathrm{T}}\mathbf{f}\,, \tag{1.2a}$$

$$\Sigma(\Pi) = \Big\{ (\hat{\sigma}(\mathbf{f}), \hat{\mu}(\mathbf{f})) \,:\, \mathbf{f} \in \Pi \Big\}\,. \tag{1.2b}$$

Examples of such Π include:

- $\mathcal{M}$ or $\mathcal{M}_+$, in which case the $\Sigma(\Pi)$ are unbounded, convex sets;
- Λ, Λ_+, Π^ℓ or Π^ℓ_+, in which case the $\Sigma(\Pi)$ are compact, nonconvex sets.

Mean-Variance Objectives (Examples in $\mathcal{M}_2$)

Similarly, when $\Pi \subset \mathcal{M}_2$ we have

$$\hat{\sigma}(\mathbf{f}) = \sqrt{\mathbf{f}^{\mathrm{T}}\mathbf{V}\,\mathbf{f}}\,,$$
$$\hat{\mu}\Big(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}\Big) = \mathbf{m}^{\mathrm{T}}\mathbf{f} + \mu_{\mathrm{si}} f^{\mathrm{si}} + \mu_{\mathrm{cl}} f^{\mathrm{cl}}\,, \tag{1.3a}$$
$$\Sigma(\Pi) = \Big\{ \Big(\hat{\sigma}(\mathbf{f})\,,\ \hat{\mu}\Big(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}\Big)\Big) \ :\ \Big(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}\Big) \in \Pi \Big\}\,. \tag{1.3b}$$

Examples of such Π include:

- $\mathcal{M}_2$, in which case the $\Sigma(\Pi)$ is an unbounded, convex set;
- Π_2^ℓ, in which case the $\Sigma(\Pi)$ is a compact, nonconvex set.

The fact that $\hat{\mu}$ is an affine function of the allocations should be clear. A proof that $\hat{\sigma}$ is a convex function of the allocations is given below.

Mean-Variance Objectives (Convexity of $\hat{\sigma}$)

Fact 1. $\hat{\sigma}(\mathbf{f})$ is a convex function over $\mathbb{R}^N$.

Proof. Let $\mathbf{f}_0$, $\mathbf{f}_1 \in \mathbb{R}^N$ with $\mathbf{f}_0 \neq \mathbf{f}_1$. The Cauchy inequality says that

$$|\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\mathbf{f}_1| \leq \sqrt{\mathbf{f}_0^{\mathrm{T}}\mathbf{V}\,\mathbf{f}_0}\sqrt{\mathbf{f}_1^{\mathrm{T}}\mathbf{V}\,\mathbf{f}_1}\,.$$

Define $\mathbf{f}_t = (1-t)\,\mathbf{f}_0 + t\,\mathbf{f}_1$ for every $t \in [0,1]$. Then by Cauchy

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

This inequality proves **Fact 1**.

Mean-Variance Objectives (Concavity of $\widehat{\Gamma}$)

Our first result about mean-variance objectives concerns their concavity. Its proof uses the facts that $\hat{\sigma}$ is convex over Π and $\hat{\mu}$ is affine over Π.

Fact 2. Let $G(\sigma, \mu)$ be a function over a convex set Σ_G in the $\sigma\mu$-plane such that

- $G(\sigma, \mu)$ is a decreasing function of σ over Σ_G, (1.4a)
- $G(\sigma, \mu)$ is concave over Σ_G. (1.4b)

Let Π be a convex set of allocations such that $\Sigma(\Pi)$ satisfies

$$\Sigma(\Pi) \subset \Sigma_G \,. \tag{1.5}$$

Then $\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu})$ given by (1.1a) is a concave function over Π.

Remark. The convexity of Σ_G and (1.5) imply that $\Sigma_G \supset \mathrm{Hull}(\Sigma(\Pi))$.

Mean-Variance Objectives (**Fact 2** Proof)

Proof. Let $(\hat{\sigma}_0, \hat{\mu}_0)$ and $(\hat{\sigma}_1, \hat{\mu}_1)$ be the values of the estimators $\hat{\sigma}$ and $\hat{\mu}$ for two distinct allocations in Π.

For every $t \in [0,1]$ let $(\hat{\sigma}_t, \hat{\mu}_t)$ be the values $\hat{\sigma}$ and $\hat{\mu}$ for the convex combination of these allocations. Because Π is convex and satisfies (1.5), we know that $(\hat{\sigma}_t, \hat{\mu}_t) \in \Sigma_G$. Because $\hat{\sigma}$ is convex over Π by **Fact 1**, while $\hat{\mu}$ is affine over Π, we have

$$\hat{\sigma}_t \leq (1-t)\,\hat{\sigma}_0 + t\,\hat{\sigma}_1\,, \qquad \hat{\mu}_t = (1-t)\,\hat{\mu}_0 + t\,\hat{\mu}_1\,.$$

Then the σ monotonicity (1.4a) followed by the concavity (1.4b) yield

$$\begin{aligned} G(\hat{\sigma}_t, \hat{\mu}_t) &\geq G\Big((1-t)\,\hat{\sigma}_0 + t\,\hat{\sigma}_1\,,\,(1-t)\,\hat{\mu}_0 + t\,\hat{\mu}_1\Big) \\ &\geq (1-t)\,G(\hat{\sigma}_0, \hat{\mu}_0) + t\,G(\hat{\sigma}_1, \hat{\mu}_1)\ , \end{aligned}$$

Therefore $\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu})$ is concave over Π. This proves **Fact 2**. □

Efficient Frontier (Introduction)

A central part of our main result about the maximization problem for $\widehat{\Gamma}$ over Π is that its maximizer must be an efficient portfolio within Π. This means that if the the efficient frontier for Π lies on the curve $\sigma = \sigma_{\mathrm{f}}(\mu)$ in the $\sigma\mu$-plane then we can introduce

$$\Gamma_{\mathrm{f}}(\mu) = G(\sigma_{\mathrm{f}}(\mu)\,,\,\mu)\,, \tag{2.6}$$

and reduce the problem of maximizing $\widehat{\Gamma}$ over Π to that of maximizing $\Gamma_{\mathrm{f}}(\mu)$ over some interval. This is a huge simplification!

- If $\sigma_{\mathrm{f}}(\mu)$ is known analytically and it and $G(\sigma, \mu)$ are sufficiently simple then an analytic solution of the problem can be found.
- If $\sigma_{\mathrm{f}}(\mu)$ is known numerically then this reduction greatly simplifies the numerical solution of the problem.

Before giving this result, we lay some groundwork about the efficient frontier.

Efficient Frontier (The Interval $\hat{\mu}(\Pi)$)

The frontier of Π is define over the set $\hat{\mu}(\Pi) \subset \mathbb{R}$ given by

$$\hat{\mu}(\Pi) = \left\{ \hat{\mu} \; : \; \text{all allocations in } \Pi \right\} . \tag{2.7}$$

Because Π is convex, it is connected. Because the continuous image of a connected set is a connected set, the facts that Π is connected and that $\hat{\mu}$ is continuous over Π imply that $\hat{\mu}(\Pi)$ is connected. But the connected subsets of $\mathbb{R}$ are the intervals, so that $\hat{\mu}(\Pi)$ is always an interval.

- If Π is $\mathcal{M}$, $\mathcal{M}_+$ or $\mathcal{M}_2$ then $\hat{\mu}(\Pi) = \mathbb{R}$.
- If $\Pi = \Lambda$ then $\hat{\mu}(\Pi) = [\mu_{\mathrm{mn}}, \mu_{\mathrm{mx}}]$.
- If $\Pi = \Pi^\ell$ for some $\ell \geq 0$ then $\hat{\mu}(\Pi) = [\mu^\ell_{\mathrm{mn}}, \mu^\ell_{\mathrm{mx}}]$, where

$$\mu^\ell_{\mathrm{mn}} = \mu_{\mathrm{mn}} - \ell \left(\mu_{\mathrm{mx}} - \mu_{\mathrm{mn}} \right) , \qquad \mu^\ell_{\mathrm{mx}} = \mu_{\mathrm{mx}} + \ell \left(\mu_{\mathrm{mx}} - \mu_{\mathrm{mn}} \right) .$$

- If Π is Λ_+, Π^ℓ_+ or Π^ℓ_2 for some $\ell \geq 0$ then $\hat{\mu}(\Pi)$ is a bounded interval that includes the risk-free rates and that can depend upon those rates.

Efficient Frontier (The Function $\sigma_{\rm f}(\mu)$)

Recall that the frontier of $\Sigma(\Pi)$ in the $\sigma\mu$-plane is given by $\sigma = \sigma_{\rm f}(\mu)$, where $\sigma_{\rm f}(\mu)$ is defined for every $\mu \in \hat{\mu}(\Pi)$ by

$$\sigma_{\rm f}(\mu) = \min\left\{\hat{\sigma} \;:\; \text{all allocations in } \Pi \text{ with } \hat{\mu} = \mu\right\}. \tag{2.8}$$

The *efficient frontier* is simply the restriction of $\sigma_{\rm f}(\mu)$ to efficient profolios. We have analytic expressions for it when Π is $\mathcal{M}$, $\mathcal{M}_+$ or $\mathcal{M}_2$.

- When $\Pi = \mathcal{M}$ then the efficient Markowitz frontier is

$$\sigma = \sigma_{\rm mf}(\mu) = \sqrt{\sigma_{\rm mv}^2 + \frac{(\mu - \mu_{\rm mv})^2}{\nu_{\rm mv}^2}} \qquad \text{for } \mu \in [\mu_{\rm mv}, \infty). \tag{2.9}$$

- When $\Pi = \mathcal{M}_+$ then the efficient Tobin frontier is

$$\sigma = \sigma_{\rm tf}(\mu) = \frac{\mu - \mu_{\rm rf}}{\nu_{\rm rf}} \qquad \text{for } \mu \in [\mu_{\rm rf}, \infty). \tag{2.10}$$

Efficient Frontier (Case $\mathcal{M}_2$)

- When $\Pi = \mathcal{M}_2$ and $\mu_{\rm mv} \leq \mu_{\rm si} < \mu_{\rm cl}$ then the efficient frontier is

$$\sigma = \sigma_{\rm f}(\mu) = \frac{\mu - \mu_{\rm si}}{\nu_{\rm si}} \qquad \text{for } \mu \in [\mu_{\rm si}, \infty)\,. \tag{2.11a}$$

- When $\Pi = \mathcal{M}_2$ and $\mu_{\rm si} < \mu_{\rm mv} \leq \mu_{\rm cl}$ then the efficient frontier is

$$\sigma = \sigma_{\rm f}(\mu) = \begin{cases} \sigma_{\rm mf}(\mu) & \text{for } \mu \in [\mu_{\rm st}, \infty)\,, \\ \frac{\mu - \mu_{\rm si}}{\nu_{\rm si}} & \text{for } \mu \in [\mu_{\rm si}, \mu_{\rm st})\,. \end{cases} \tag{2.11b}$$

- When $\Pi = \mathcal{M}_2$ and $\mu_{\rm si} < \mu_{\rm cl} < \mu_{\rm mv}$ then the efficient frontier is

$$\sigma = \sigma_{\rm f}(\mu) = \begin{cases} \frac{\mu - \mu_{\rm cl}}{\nu_{\rm cl}} & \text{for } \mu \in [\mu_{\rm ct}, \infty)\,, \\ \sigma_{\rm mf}(\mu) & \text{for } \mu \in [\mu_{\rm st}, \mu_{\rm ct})\,, \\ \frac{\mu - \mu_{\rm si}}{\nu_{\rm si}} & \text{for } \mu \in [\mu_{\rm si}, \mu_{\rm st})\,. \end{cases} \tag{2.11c}$$

Efficient Frontier (General Case)

In general the function $\sigma_{\mathrm{f}}(\mu)$ has been approximated numerically at select points in $\hat{\mu}(\Pi)$ and is interpolated at other points in $\hat{\mu}(\Pi)$.

- If risk-free assets are excluded then $\Pi \subset \mathcal{M}$ and the efficient frontier restricts $\sigma_{\mathrm{f}}(\mu)$ to the interval $\hat{\mu}(\Pi) \cap [\mu_{\mathrm{mv}}^{\mathrm{f}}, \infty)$, where $\mu_{\mathrm{mv}}^{\mathrm{f}}$ is the minimizer of $\sigma_{\mathrm{f}}(\mu)$.
- If risk-free assets are included with the one-rate model then $\Pi \subset \mathcal{M}_+$ and the efficient frontier restricts $\sigma_{\mathrm{f}}(\mu)$ to the interval $\hat{\mu}(\Pi) \cap [\mu_{\mathrm{rf}}, \infty)$.
  - If $\Pi = \Lambda_+$ and $\mu_{\mathrm{rf}} < \mu_{\mathrm{mx}}$ then the interval is $[\mu_{\mathrm{rf}}, \mu_{\mathrm{mx}}]$.
  - If $\Pi = \Pi_+^{\ell}$ and $\mu_{\mathrm{mn}} < \mu_{\mathrm{rf}} < \mu_{\mathrm{mx}}$ then the interval is $[\mu_{\mathrm{rf}}, \mu_{\mathrm{mx}}^{\ell}]$.
- If risk-free assets are included with the two-rate model then $\Pi \subset \mathcal{M}_2$ and the efficient frontier restricts $\sigma_{\mathrm{f}}(\mu)$ to the interval $\hat{\mu}(\Pi) \cap [\mu_{\mathrm{si}}, \infty)$.
  - If $\Pi = \Pi_2^{\ell}$ and $\mu_{\mathrm{mn}} < \mu_{\mathrm{cl}}$ and $\mu_{\mathrm{si}} < \mu_{\mathrm{mx}}$ then the interval is $[\mu_{\mathrm{si}}, \mu_{\mathrm{mx}}^{\ell}]$.

Below we prove that $\sigma_{\mathrm{f}}(\mu)$ is always convex over $\hat{\mu}(\Pi)$. More cannot be expected because the Tobin frontier is not strictly convex.

Efficient Frontier (Convexity of $\sigma_{\rm f}(\mu)$)

Fact 3. The function $\sigma_{\rm f}(\mu)$ is convex over $\hat{\mu}(\Pi)$.

Remark. We give the proof for $\Pi \subset \mathcal{M}_+$. The rest is left as an exercise.

Proof. Let μ_0 and $\mu_1 \in \hat{\mu}(\Pi)$ with $\mu_0 < \mu_1$. Let $\mathbf{f}_0 \in \Pi$ with $\hat{\mu}(\mathbf{f}_0) = \mu_0$ and $\mathbf{f}_1 \in \Pi$ with $\hat{\mu}(\mathbf{f}_1) = \mu_1$ be arbitrary. Fix $t \in [0,1]$ and set

$$\mu_t = (1-t)\,\mu_0 + t\,\mu_1\,, \qquad \mathbf{f}_t = (1-t)\,\mathbf{f}_0 + t\,\mathbf{f}_1\,.$$

Because Π is convex and $\hat{\mu}(\mathbf{f})$ is affine, we know $\mathbf{f}_t \in \Pi$ and $\hat{\mu}(\mathbf{f}_t) = \mu_t$. Then definition (2.8) of $\sigma_{\rm f}(\mu)$ and the convexity of $\hat{\sigma}(\mathbf{f})$ show

$$\sigma_{\rm f}(\mu_t) \leq \hat{\sigma}(\mathbf{f}_t) \leq (1-t)\,\hat{\sigma}(\mathbf{f}_0) + t\,\hat{\sigma}(\mathbf{f}_1)\,.$$

Minimizing the right-hand side over the arbitrary $\mathbf{f}_0$ and $\mathbf{f}_1$, we obtain

$$\sigma_{\rm f}(\mu_t) \leq (1-t)\,\sigma_{\rm f}(\mu_0) + t\,\sigma_{\rm f}(\mu_1)\,.$$

But $t \in [0,1]$ was arbitrary. Therefore **Fact 3** is proved. □

Efficient Frontier (Maximizers are Frontier

Our first result abour maximizers says that they are frontier portfolios.

Fact 4. Let $G(\sigma,\mu)$ be a function over a convex set Σ_G in the $\sigma\mu$-plane such that

- $G(\sigma,\mu)$ is a decreasing function of σ over Σ_G . (2.12)

Let Π be a convex set of allocations such that $\Sigma(\Pi)$ satisfies

$$\Sigma(\Pi) \subset \Sigma_G \,. \tag{2.13}$$

Any maximizer of $\widehat{\Gamma} = G(\hat{\sigma},\hat{\mu})$ over Π must be a frontier portfolio of Π.

Proof. Any allocation that is not a frontier portfolio of Π must satisfy $\sigma_{\mathrm{f}}(\hat{\mu}) < \hat{\sigma}$. The monotonicity condition (2.12) then implies that

$$G(\sigma_{\mathrm{f}}(\hat{\mu})\,,\,\hat{\mu}) > G(\hat{\sigma}\,,\,\hat{\mu})\;,$$

whereby $\widehat{\Gamma}$ is larger for the frontier portfolio associated with $(\sigma_{\mathrm{f}}(\hat{\mu})\,,\,\hat{\mu})$. □

Efficient Frontier (Efficiency Monotonicity)

In our next result we replace the σ monotonicity condition (2.12) with a stronger condition. Given any two points (σ_0, μ_0) and (σ_1, μ_1) in the $\sigma\mu$-plane, we say that (σ_1, μ_1) is *more efficient* than (σ_0, μ_0), denoted $(\sigma_1, \mu_1) \succ (\sigma_0, \mu_0)$, when

$$\sigma_1 \leq \sigma_0 \,, \qquad \mu_1 \geq \mu_0 \,, \qquad (\sigma_1, \mu_1) \neq (\sigma_0, \mu_0) \,. \tag{2.14}$$

Of course, this notion coincides with that of Markowitz efficiency when the points represent the volatilities and return means of portfolios.

Definiton 1. We say that $G(\sigma, \mu)$ *increases with efficiency* over a subset Σ of the $\sigma\mu$-plane when for every (σ_0, μ_0), $(\sigma_1, \mu_1) \in \Sigma$ we have

$$(\sigma_1, \mu_1) \succ (\sigma_0, \mu_0) \quad \Longrightarrow \quad G(\sigma_1, \mu_1) > G(\sigma_0, \mu_0) \,. \tag{2.15}$$

Remark. Because $(\sigma_1, \mu) \succ (\sigma_0, \mu)$ if and only if $\sigma_1 < \sigma_0$, we see that if (2.15) holds over Σ them $G(\sigma, \mu)$ is a decreasing function of σ over Σ.

Efficient Frontier (Maximizers are Efficient)

We now replace the σ monotonicity condition (2.12) in **Fact 4** with an efficiency monotonicity condition as defined by (2.15). This will allow us to conclude that maximizers are efficient.

Fact 5. Let $G(\sigma, \mu)$ be a function over a convex set Σ_G in the $\sigma\mu$-plane such that

- $G(\sigma, \mu)$ increases with efficiency over Σ_G. (2.16)

Let Π be a convex set of allocations such that $\Sigma(\Pi)$ satisfies

$$\Sigma(\Pi) \subset \Sigma_G \,. \tag{2.17}$$

Any maximizer of $\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu})$ over Π must be an efficient frontier portfolio of Π.

Maximization Problem (**Fact 5** Proof)

Proof of Fact 5. Let $\widehat{\Gamma}_{\mathrm{mx}}$ be the maximum of $\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu})$ over Π. Then **Fact 4** says that any maximizer over Π is a frontier portfolio. Let $(\hat{\sigma}_0, \hat{\mu}_0) \in \Sigma(\Pi)$ be the values of $\hat{\sigma}$ and $\hat{\mu}$ at such a maximizer.

If the maximizer is not efficient in Π then there exists another allocation in Π at which $\hat{\sigma}$ and $\hat{\mu}$ have values $(\hat{\sigma}_1, \hat{\mu}_1) \in \Sigma(\Pi)$ such that

$$(\hat{\sigma}_1, \hat{\mu}_1) \succ (\hat{\sigma}_0, \hat{\mu}_0) \,.$$

Because $\Sigma(\Pi) \subset \Sigma_G$ by (2.17), we see from the efficiency monotonicity (2.16) that

$$G(\hat{\sigma}_1, \hat{\mu}_1) > G(\hat{\sigma}_0, \hat{\mu}_0) = \widehat{\Gamma}_{\mathrm{mx}} \,.$$

This contradicts the fact that $\widehat{\Gamma}_{\mathrm{mx}}$ is the maximum of $\widehat{\Gamma}$ over Π. Therefore the maximizer must be efficient. This proves **Fact 5**. □

Level Sets and Convexity (Concavity of G)

The uniqueness of the maximizer will require two additional hypotheses.

The first hypothesis is that $G(\sigma, \mu)$ is concave over the convex set Σ_G. This means that for every (σ_0, μ_0), $(\sigma_1, \mu_1) \in \Sigma_G$ and every $t \in [0, 1]$ we have

$$
\begin{aligned}
&G\Big((1-t)\,\sigma_0 + t\,\sigma_1\,,\,(1-t)\,\mu_0 + t\,\mu_1\Big) \\
&\geq (1-t)\,G(\sigma_0, \mu_0) + t\,G(\sigma_1, \mu_1)\,.
\end{aligned}
$$

This insures that for every $\Gamma \in \mathbb{R}$

$$
\text{the set} \quad \Big\{(\sigma, \mu) \in \Sigma_G \,:\, G(\sigma, \mu) \geq \Gamma\Big\} \quad \text{is convex}\,. \tag{3.18}
$$

This set in nonempty if and only if Γ is in the range of G over Σ_G.

Level Sets and Convexity (Level Sets)

The boundary of the set (3.18) is the *level set*

$$\Big\{(\sigma,\mu)\in\Sigma_G\,:\,G(\sigma,\mu)=\Gamma\Big\}. \tag{3.19}$$

If $G(\sigma,\mu)$ is twice continuously differentiable over Σ_G and its gradient never vanishes over Σ_G then the *Implicit Function Theorem* says that every level set is the union twice continuously differentiable curves in Σ_G.

Definition 2. We say that these level set curves are *curved* if they have nonzero curvature at every point.

The second hypothesis is that the level set curves are curved. Below we derive conditions that imply this hypothesis. We will assume $G(\sigma,\mu)$ is twice continuously differentiable and denote its partial derivatives with subscripts.

Level Sets and Convexity (Curved Level Sets)

If $G_\mu(\sigma,\mu) > 0$ over Σ_G then the level curve associated with Γ can be parameterized by σ. Let $\mu = \mu^\Gamma(\sigma)$ be the unique solution of

$$G(\sigma,\mu) = \Gamma\,. \tag{3.20}$$

By taking the derivative of (3.20) with respect to σ we find

$$G_\sigma(\sigma,\mu) + G_\mu(\sigma,\mu)\,\frac{\partial\mu}{\partial\sigma} = 0\,.$$

Because $G_\mu(\sigma,\mu) > 0$, this can be solved to obtain

$$\frac{\partial\mu}{\partial\sigma} = -\frac{G_\sigma(\sigma,\mu)}{G_\mu(\sigma,\mu)}\,. \tag{3.21a}$$

By taking the second derivative of (3.20) with respect to σ we find

$$\frac{\partial^2\mu}{\partial\sigma^2} = -\frac{1}{G_\mu^3}\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix}\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix}\begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix}. \tag{3.21b}$$

Level Sets and Convexity (Curved Level Sets)

Alternatively, if $G_\sigma(\sigma,\mu) < 0$ over Σ_G then the level curve associated with Γ can be parameterized by μ. Let $\sigma = \sigma^\Gamma(\mu)$ be the unique solution of (3.20). By taking the derivative of (3.20) with respect to μ we find

$$G_\sigma(\sigma,\mu)\,\frac{\partial\sigma}{\partial\mu} + G_\mu(\sigma,\mu) = 0\,.$$

Because $G_\sigma(\sigma,\mu) < 0$, this can be solved to obtain

$$\frac{\partial\sigma}{\partial\mu} = -\frac{G_\mu(\sigma,\mu)}{G_\sigma(\sigma,\mu)}\,. \tag{3.22a}$$

By taking the second derivative of with respect to μ we find

$$\frac{\partial^2\sigma}{\partial\mu^2} = -\frac{1}{G_\sigma^3}\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix}\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix}\begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix}\,. \tag{3.22b}$$

Level Sets and Convexity (Curved Level Sets)

The hypothesis that these level set curves are curved is satisfied when either

$$\frac{\partial^2 \mu^\Gamma}{\partial \sigma^2} > 0 \qquad \text{or} \qquad \frac{\partial^2 \sigma^\Gamma}{\partial \mu^2} < 0\,. \tag{3.23a}$$

It is clear from (3.21b) and (3.22b) that this

$$\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix} \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix} < 0 \quad \text{over } \Sigma_G\,. \tag{3.23b}$$

The hypothesis that $G(\sigma, \mu)$ is concave over Σ_G and the hypothesis that these level set curves are curved are both satisfied when

$$\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \quad \text{is negative definite over } \Sigma_G\,. \tag{3.24}$$

Level Sets and Convexity (Uniqueness)

Our main result says the maximizer is also unique.

Fact 6. Let $G(\sigma, \mu)$ be a function over a convex set Σ_G in the $\sigma\mu$-plane such that

- $G(\sigma, \mu)$ is a decreasing function of σ over Σ_G , (3.25a)
- $G(\sigma, \mu)$ is concave over Σ_G , (3.25b)
- $G(\sigma, \mu)$ has curved level sets in Σ_G . (3.25c)

Let Π be a convex set of allocations such that $\Sigma(\Pi)$ satisfies

$$\Sigma(\Pi) \subset \Sigma_G \,. \tag{3.26}$$

Any maximizer of $\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu})$ over Π must be an efficient frontier portfolio of Π and is unique.

Maximization Problem (**Fact 6** Proof)

Proof of Fact 6. Suppose that the maximum of $\widehat{\Gamma}$ over Π is $\widehat{\Gamma}_{\mathrm{mx}}$ and that there are two maximizers. At these maximizers let $\hat{\sigma}$ and $\hat{\mu}$ have values

$$(\hat{\sigma}_0 \, , \, \hat{\mu}_0) \, , \qquad (\hat{\sigma}_1 \, , \, \hat{\mu}_1) \, . \tag{3.27}$$

By **Fact 4** these maximizers must be frontier portfolios. Because there is a unique frontier portfolio for each $\mu \in \hat{\mu}(\Pi)$, we see that $\hat{\mu}_0 \neq \hat{\mu}_1$. Therefore the points in $\Sigma(\Pi)$ given by (3.27) are distinct.

For every $t \in (0,1)$ let $(\hat{\sigma}_t, \hat{\mu}_t)$ be the values $\hat{\sigma}$ and $\hat{\mu}$ for the convex combination of these allocations. Because Π is convex and satisfies (3.26), we know that $(\hat{\sigma}_t, \hat{\mu}_t) \in \Sigma_G$. Because $\hat{\sigma}$ is convex over Π by **Fact 1**, while $\hat{\mu}$ is affine over Π, we have

$$\hat{\sigma}_t \leq (1-t)\,\hat{\sigma}_0 + t\,\hat{\sigma}_1 \, , \qquad \hat{\mu}_t = (1-t)\,\hat{\mu}_0 + t\,\hat{\mu}_1 \, . \tag{3.28}$$

Maximization Problem (**Fact 6** Proof)

The σ monotonicity (3.25a) and (3.28) followed by the combination of

- the fact the points in $\Sigma(\Pi)$ given by (3.27) are distinct,
- the fact $\Sigma(\Pi) \subset \Sigma_G$ by (3.26),
- the concavity (3.25b) of $G(\sigma, \mu)$ over Σ_G,

then yield

$$\begin{aligned}\widehat{\Gamma}_{\mathrm{mx}} \geq G(\hat{\sigma}_t, \hat{\mu}_t) &\geq G\Big((1-t)\,\hat{\sigma}_0 + t\,\hat{\sigma}_1\,,\,(1-t)\,\hat{\mu}_0 + t\,\hat{\mu}_1\Big) \\ &\geq (1-t)\,G(\hat{\sigma}_0, \hat{\mu}_0) + t\,G(\hat{\sigma}_1, \hat{\mu}_1) = \widehat{\Gamma}_{\mathrm{mx}}\,.\end{aligned}$$

But this says the line segment connecting the points in Σ_G given by (3.27) is a level set, which contradicts (3.25c). Therefore there cannot be two maximizers. This proves **Fact 6**. □

Applications (Introduction)

In order to apply either **Fact 4**, **Fact 5** or **Fact 6** to a mean-variance objective $\widehat{\Gamma}$ in the form

$$\widehat{\Gamma} = G(\hat{\sigma}, \hat{\mu}) \ , \tag{4.29}$$

we must

1. identify a convex subset Σ_G over which $G(\sigma, \mu)$ satisfies the hypotheses in each fact,
2. identify convex sets of allocations Π that satisfy $\Sigma(\Pi) \subset \Sigma_G$.

Here we will try to do this for the mean-variance estimators derived earlier. We will see that this program can be completed for most of those estimators, but not all. The ones where it fails to complete breakdowwn at the first step. Later we will learn from these cases how the troublesome hypotheses can be weakened without weakening the conclusions.

Applications (Hypotheses)

The hypotheses on the convex set Σ_G that appear in either **Fact 4**, **Fact 5** or **Fact 6** are

- $G(\sigma, \mu)$ is a decreasing function of σ over Σ_G , (4.30a)
- $G(\sigma, \mu)$ increases with efficiency over Σ_G , (4.30b)
- $G(\sigma, \mu)$ is concave over Σ_G , (4.30c)
- $G(\sigma, \mu)$ has curved level sets in Σ_G . (4.30d)

The $G(\sigma, \mu)$ for the mean-variance objectives derived earlier are all smooth over their natural domains, so the above hypotheses can be verified by taking partial derivatves.

Applications (Partial Derivative Tests)

For example:

- hypothesis (4.30a) holds over sets where $G_\sigma < 0$;
- hypothesis (4.30b) holds over sets where $G_\sigma < 0$ and $G_\mu > 0$;
- hypothesis (4.30c) holds over sets where

$$\text{the Hessian} \quad \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \quad \text{is nonpositive definite}\,; \qquad (4.31a)$$

- hypothesis (4.30d) holds over sets where the Hessian satisfies

$$\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix} \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix} < 0\,; \qquad (4.31b)$$

- hypotheses (4.30c) and (4.30d) both hold over sets where the Hessian is negative definite.

Applications (Some Examples)

If $\widehat{\Gamma}$ is $\widehat{\Gamma}^{\chi}_{\mathrm{p}}$, $\widehat{\Gamma}^{\chi}_{\mathrm{q}}$, $\widehat{\Gamma}^{\chi}_{\mathrm{r}}$, $\widehat{\Gamma}^{\chi}_{\mathrm{s}}$, $\widehat{\Gamma}^{\chi}_{\mathrm{t}}$, or $\widehat{\Gamma}^{\chi}_{\mathrm{u}}$ for some $\chi \geq 0$ then

$$G^{\chi}_{\mathrm{p}}(\sigma,\mu) = \mu - \tfrac{1}{2}\sigma^2 - \chi\,\sigma\,, \tag{4.32a}$$

$$G^{\chi}_{\mathrm{q}}(\sigma,\mu) = \mu - \tfrac{1}{2}\mu^2 - \tfrac{1}{2}\sigma^2 - \chi\,\sigma\,, \tag{4.32b}$$

$$G^{\chi}_{\mathrm{r}}(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\sigma^2 - \chi\,\sigma\,, \tag{4.32c}$$

$$G^{\chi}_{\mathrm{s}}(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{1+\mu} - \chi\,\sigma\,, \tag{4.32d}$$

$$G^{\chi}_{\mathrm{t}}(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{(1+\mu)^2} - \chi\,\sigma\,, \tag{4.32e}$$

$$G^{\chi}_{\mathrm{u}}(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{(1+\mu)^2} - \chi\,\frac{\sigma}{1+\mu}\,. \tag{4.32f}$$

Applications (Natural Domains)

These are the parabolic, quadratic, reasonable, sensible, Taylor, and ultimate estimators respectively. Their respective natural domains are

$$\Sigma_{\mathrm{p}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0\} \, , \tag{4.33a}$$

$$\Sigma_{\mathrm{q}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0\} \, , \tag{4.33b}$$

$$\Sigma_{\mathrm{r}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0 \, , \, 1 + \mu > 0\} \, , \tag{4.33c}$$

$$\Sigma_{\mathrm{s}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0 \, , \, 1 + \mu > 0\} \, , \tag{4.33d}$$

$$\Sigma_{\mathrm{t}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0 \, , \, 1 + \mu > 0\} \, , \tag{4.33e}$$

$$\Sigma_{\mathrm{u}} = \{(\sigma, \mu) \in \mathbb{R}^2 \, : \, \sigma \geq 0 \, , \, 1 + \mu > 0\} \, . \tag{4.33f}$$

These natural domains are convex subsets of $\mathbb{R}^2$ that satisfy

$$\Sigma_{\mathrm{p}} = \Sigma_{\mathrm{q}} \supset \Sigma_{\mathrm{r}} = \Sigma_{\mathrm{s}} = \Sigma_{\mathrm{t}} = \Sigma_{\mathrm{u}} \, .$$

Our first goal is to identify subsets of these domains that can play the role of Σ_G in the hypotheses (4.30).

Applications (Parabolic)

For the *parabolic estimator* we see from (4.32a) that

$$G(\sigma, \mu) = \mu - \tfrac{1}{2}\,\sigma^2 - \chi\,\sigma\,,$$

and from (4.33a) that

$$\Sigma_{\mathrm{p}} = \Big\{(\sigma, \mu) \in \mathbb{R}^2 \,:\, \sigma \geq 0\Big\}\,.$$

Taking partial derivatives we find that

$$\begin{gathered} G_\sigma = -\sigma - \chi\,, \qquad G_\mu = 1\,, \\ \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}\,. \end{gathered} \tag{4.34}$$

Applications (Parabolic)

We see from (4.34) that for every $\chi \geq 0$

- $G(\sigma, \mu)$ increases with efficiency over Σ_{p};
- $G(\sigma, \mu)$ is convex over Σ_{p}, but it is not strictly convex over any subset of Σ_{p};
- $G(\sigma, \mu)$ has curved level sets in Σ_G because

$$
\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix} \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix}
= \begin{pmatrix} 1 & \sigma + \chi \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \sigma + \chi \end{pmatrix} = -1 < 0 \, .
$$

Therefore we can apply either **Fact 4**, **Fact 5** or **Fact 6** with $\Sigma_G = \Sigma_{\mathrm{p}}$.

Applications (Quadratic)

For the *quadratic estimator* we see from (4.32b) that

$$G(\sigma,\mu) = \mu - \tfrac{1}{2}\,\mu^2 - \tfrac{1}{2}\,\sigma^2 - \chi\,\sigma\,,$$

and from (4.33b) that

$$\Sigma_{\mathrm{q}} = \left\{(\sigma,\mu)\in\mathbb{R}^2 \,:\, \sigma \geq 0\right\}.$$

Taking partial derivatives we find that

$$\begin{gathered} G_\sigma = -\sigma - \chi\,, \qquad G_\mu = 1 - \mu\,, \\ \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}. \end{gathered} \tag{4.35}$$

Applications (Quadratic)

We see from (4.35) that for every $\chi \geq 0$

- $G(\sigma, \mu)$ is a decreasing function of σ over $\Sigma_{\rm q}$,
- $G(\sigma, \mu)$ increases with efficiency over the subset of $\Sigma_{\rm q}$ where $\mu \leq 1$,
- $G(\sigma, \mu)$ is strictly convex over $\Sigma_{\rm q}$.

Therefore we can apply either **Fact 4** or **Fact 6** with $\Sigma_G = \Sigma_{\rm q}$, and can apply **Fact 5** with

$$\Sigma_G = \Big\{ (\sigma, \mu) \in \Sigma_{\rm q} \ : \ \mu \leq 1 \Big\} .$$

This suggests that when $\Pi \subset \mathcal{M}_+$ it should satisfy $\Pi \subset \Omega_{\rm q}$, where

$$\Omega_{\rm r} = \Big\{ \mathbf{f} \in \mathcal{M}_+ \ : \ \hat{\mu}(\mathbf{f}) \leq 1 \Big\} ,$$

Applications (Reasonable)

For the *reasonable estimator* we see from (4.32c) that

$$G(\sigma, \mu) = \log(1 + \mu) - \tfrac{1}{2}\,\sigma^2 - \chi\,\sigma\,,$$

and from (4.33c) that

$$\Sigma_{\mathrm{r}} = \Big\{(\sigma, \mu) \in \mathbb{R}^2 \,:\, \sigma \geq 0\,,\, 1 + \mu > 0\Big\}\,.$$

Taking partial derivatives we find that

$$\begin{gathered} G_\sigma = -\sigma - \chi\,, \qquad G_\mu = \tfrac{1}{1+\mu}\,, \\ \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{(1+\mu)^2} \end{pmatrix}\,. \end{gathered} \tag{4.36}$$

Applications (Reasonable)

We see from (4.36) that for every $\chi \geq 0$

- $G(\sigma, \mu)$ increases with efficiency over Σ_{r},
- $G(\sigma, \mu)$ is strictly convex over Σ_{r}.

Therefore we can apply either **Fact 4**, **Fact 5** or **Fact 6** with $\Sigma_G = \Sigma_{\mathrm{r}}$.

This suggests that when $\Pi \subset \mathcal{M}_+$ it should satisfy $\Pi \subset \Omega_{\mathrm{r}}$, where

$$\Omega_{\mathrm{r}} = \left\{ \mathbf{f} \in \mathcal{M}_+ \;:\; 1 + \hat{\mu}(\mathbf{f}) > 0 \right\},$$

and when $\Pi \subset \mathcal{M}_2$ it should satisfy $\Pi \subset \Omega_{\mathrm{r}}$, where

$$\Omega_{\mathrm{r}} = \left\{ (\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) \in \mathcal{M}_2 \;:\; 1 + \hat{\mu}(\mathbf{f}, f^{\mathrm{si}}, f^{\mathrm{cl}}) > 0 \right\}.$$

Applications (Sensible)

For the *sensible estimator* we see from (4.32d) that

$$G(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{1+\mu} - \chi\,\sigma\,,$$

and from (4.33d) that

$$\Sigma_{\mathrm{s}} = \Big\{(\sigma,\mu)\in\mathbb{R}^2\,:\,\sigma\geq 0\,,\,1+\mu>0\Big\}\,.$$

Taking partial derivatives we find that

$$\begin{gathered} G_\sigma = -\tfrac{\sigma}{1+\mu} - \chi\,, \qquad G_\mu = \tfrac{1}{1+\mu} + \tfrac{1}{2}\,\tfrac{\sigma^2}{(1+\mu)^2}\,, \\ \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -\tfrac{1}{1+\mu} & \tfrac{\sigma}{(1+\mu)^2} \\ \tfrac{\sigma}{(1+\mu)^2} & -\tfrac{1}{(1+\mu)^2} - \tfrac{\sigma^2}{(1+\mu)^3} \end{pmatrix}. \end{gathered} \tag{4.37}$$

Applications (Sensible)

We see from (4.37) that

$$\det\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \frac{1}{(1+\mu)^3}$$

and that for every $\chi \geq 0$

- $G(\sigma,\mu)$ increases with efficiency over Σ_{s},
- $G(\sigma,\mu)$ is strictly convex over Σ_{s}.

Therefore we can apply either **Fact 4**, **Fact 5** or **Fact 6** with $\Sigma_G = \Sigma_{\mathrm{s}}$.

This suggests that when $\Pi \subset \mathcal{M}_+$ it should satisfy $\Pi \subset \Omega_{\mathrm{s}}$, where

$$\Omega_{\mathrm{s}} = \left\{ \mathbf{f} \in \mathcal{M}_+ \ : \ 1 + \hat{\mu}(\mathbf{f}) > 0 \right\}.$$

Applications (Taylor)

For the *Taylor estimator* we see from (4.32e) that

$$G(\sigma, \mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{(1+\mu)^2} - \chi\,\sigma\,,$$

and from (4.33e) that

$$\Sigma_{\rm t} = \Big\{(\sigma, \mu) \in \mathbb{R}^2 \,:\, \sigma \geq 0\,,\; 1+\mu > 0\Big\}\,.$$

Taking partial derivatives we find that

$$\begin{gathered} G_\sigma = -\tfrac{\sigma}{(1+\mu)^2} - \chi\,, \qquad G_\mu = \tfrac{1}{1+\mu} + \tfrac{\sigma^2}{(1+\mu)^3}\,, \\ \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -\frac{1}{(1+\mu)^2} & \frac{2\sigma}{(1+\mu)^3} \\ \frac{2\sigma}{(1+\mu)^3} & -\frac{1}{(1+\mu)^2} - \frac{3\sigma^2}{(1+\mu)^4} \end{pmatrix}\,. \end{gathered} \tag{4.38}$$

Applications (Taylor)

We see from (4.38) that

$$\det\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \frac{1}{(1+\mu)^4}\left(1 - \frac{\sigma^2}{(1+\mu)^2}\right),$$

and that for every $\chi \geq 0$

- $G(\sigma, \mu)$ increases with efficiency over Σ_{t},
- $G(\sigma, \mu)$ is strictly convex over the subset of Σ_{t} where $1 + \mu > \sigma$.

Therefore we can apply either **Fact 4** or **Fact 5** with $\Sigma_G = \Sigma_{\mathrm{t}}$, and can apply **Fact 6** with

$$\Sigma_G = \left\{(\sigma, \mu) \in \Sigma_{\mathrm{t}} \,:\, 1 + \mu > \sigma\right\}.$$

This suggests that when $\Pi \subset \mathcal{M}_+$ it should satisfy $\Pi \subset \Omega_{\mathrm{t}}$, where

$$\Omega_{\mathrm{t}} = \left\{\mathbf{f} \in \mathcal{M}_+ \,:\, 1 + \hat{\mu}(\mathbf{f}) > \hat{\sigma}(\mathbf{f})\right\}.$$

Applications (Taylor)

When $1 + \mu = \sigma$ the partial derivatives (4.38) become

$$G_\sigma = -\frac{1}{1+\mu} - \chi\,, \qquad G_\mu = \frac{2}{1+\mu}\,,$$
$$\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = -\frac{1}{(1+\mu)^2} \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}\,,$$

whereby

$$\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix} \begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} \begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix} = -\frac{4\,\chi^2}{(1+\mu)^2}\,.$$

So the curved level set hypothesis holds for $\chi > 0$, but not for $\chi = 0$.

Applications (Ultimate)

For the *ultimate estimator* we see from (4.32f) that

$$G(\sigma,\mu) = \log(1+\mu) - \tfrac{1}{2}\,\frac{\sigma^2}{(1+\mu)^2} - \chi\,\frac{\sigma}{1+\mu}\,,$$

and from (4.33f) that

$$\Sigma_{\mathrm{u}} = \Big\{(\sigma,\mu)\in\mathbb{R}^2 \,:\, \sigma \geq 0\,,\; 1+\mu > 0\Big\}\,.$$

Taking partial derivatives we find that

$$\begin{gathered}
G_\sigma = -\tfrac{\sigma}{(1+\mu)^2} - \tfrac{\chi}{1+\mu}\,, \qquad G_\mu = \tfrac{1}{1+\mu} + \tfrac{\sigma^2}{(1+\mu)^3} + \tfrac{\chi\,\sigma}{(1+\mu)^2}\,, \\
\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \begin{pmatrix} -\tfrac{1}{(1+\mu)^2} & \tfrac{2\sigma}{(1+\mu)^3} + \tfrac{\chi}{(1+\mu)^2} \\ \tfrac{2\sigma}{(1+\mu)^3} + \tfrac{\chi}{(1+\mu)^2} & -\tfrac{1}{(1+\mu)^2} - \tfrac{3\sigma^2}{(1+\mu)^4} - \tfrac{2\chi\,\sigma}{(1+\mu)^3} \end{pmatrix}.
\end{gathered} \tag{4.39}$$

Applications (Ultimate)

We see from (4.39) that

$$\det\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = \frac{1}{(1+\mu)^4}\left(1 - \left(\frac{\sigma}{1+\mu} + \chi\right)^2\right),$$

and that for every $\chi \in [0,1)$

- $G(\sigma,\mu)$ increases with efficiency over Σ_{u},
- $G(\sigma,\mu)$ is strictly convex over the subset of Σ_{u} where $1+\mu > \frac{\sigma}{1-\chi}$.

Therefore we can apply either **Fact 4** or **Fact 5** with $\Sigma_G = \Sigma_{\mathrm{u}}$, and can apply **Fact 6** with

$$\Sigma_G = \left\{(\sigma,\mu) \in \Sigma_{\mathrm{u}} \,:\, (1-\chi)\,(1+\mu) > \sigma\right\}.$$

This suggests that when $\Pi \subset \mathcal{M}_+$ it should satisfy $\Pi \subset \Omega_{\mathrm{u}}^{\chi}$, where

$$\Omega_{\mathrm{t}}^{\chi} = \left\{\mathbf{f} \in \mathcal{M}_+ \,:\, (1-\chi)\,(1+\hat{\mu}(\mathbf{f})) > \hat{\sigma}(\mathbf{f})\right\}.$$

Applications (Ultimate)

When $(1-\chi)(1+\mu)=\sigma$ the partial derivatives (4.39) become

$$G_\sigma = -\frac{1}{1+\mu}\,, \qquad G_\mu = \frac{2-\chi}{1+\mu}\,,$$

$$\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix} = -\frac{1}{(1+\mu)^2}\begin{pmatrix} 1 & -(2-\chi) \\ -(2-\chi) & (2-\chi)^2 \end{pmatrix}\,,$$

whereby

$$\begin{pmatrix} G_\mu & -G_\sigma \end{pmatrix}\begin{pmatrix} G_{\sigma\sigma} & G_{\sigma\mu} \\ G_{\mu\sigma} & G_{\mu\mu} \end{pmatrix}\begin{pmatrix} G_\mu \\ -G_\sigma \end{pmatrix} = 0\,.$$

So the curved level set hypothesis does not hold for any $\chi \in [0,1)$.