Financial Engineering: Portfolio Optimization

Daniel P. Palomar

Hong Kong University of Science and Technology (HKUST)

ELEC5470 - Convex Optimization

Fall 2017-18, HKUST, Hong Kong

Outline of Lecture

- Data model
- Return, risk measures, and Sharpe ratio
- Basic mean-variance portfolio formulation
- Max Sharpe ratio portfolio formulation
- CVaR portfolio formulation
- Robust portfolio design
- Index replication with ℓ_1-norm minimization

Data Model

- Consider M financial assets, e.g., stocks of a major index such as S&P 500 and Hang Seng Index.

- Denote the absolute prices at time t by $p_{m,t}$, $m = 1, \ldots, M$.

- The simple (linear) and continuously-compounded (logarithmic) returns are

$$\begin{aligned} r_{m,t} &= \left(p_{m,t} - p_{m,t-1}\right) / p_{m,t-1} \\ y_{m,t} &= \log\left(p_{m,t}\right) - \log\left(p_{m,t-1}\right). \end{aligned}$$

- Observe that, for small values of the return, we have $y_{m,t} \approx r_{m,t}$.

- We will mainly use linear returns, stacked as $\mathbf{r}_t \in \mathbb{R}^M$, and assume that they are i.i.d. random variables with mean $\boldsymbol{\mu} \in \mathbb{R}^M$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{M \times M}$.

Return

- Suppose we invest 1\$ in the mth asset at time $t-1$, then at time t we will have an amount of 1\$$\times$ $(p_{m,t}/p_{m,t-1})$.

- The relative benefit or return will be $(p_{m,t}/p_{m,t-1} - 1)$ which is precisely equal to $r_{m,t}$.

- If now we invest 1\$ distributed over the M assets according to the portfolio weights $\mathbf{w}$ (with $\mathbf{1}^T\mathbf{w} = 1$), then the overall return will be $\mathbf{w}^T\mathbf{r}_t$.

- The key quantity is then the **random return** $\mathbf{w}^T\mathbf{r}_t$.

- The **expected** or **mean return** is $\mathbf{w}^T\boldsymbol{\mu}$ and the **variance** is $\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$.

Risk Control

- In finance, the mean return is very relevant as it quantifies the average benefit.

- However, in practice, the average performance is not good enough and one needs to control the probability of going bankrupt.

- Risk measures control how risky an investment strategy is.

- The most basic measure of risk is given by the **variance**: a higher variance means that there are large peaks in the distribution which may cause a big loss.

- There are more sophisticated risk measures such as Value-at-Risk (VaR) and Conditional VaR (CVaR).

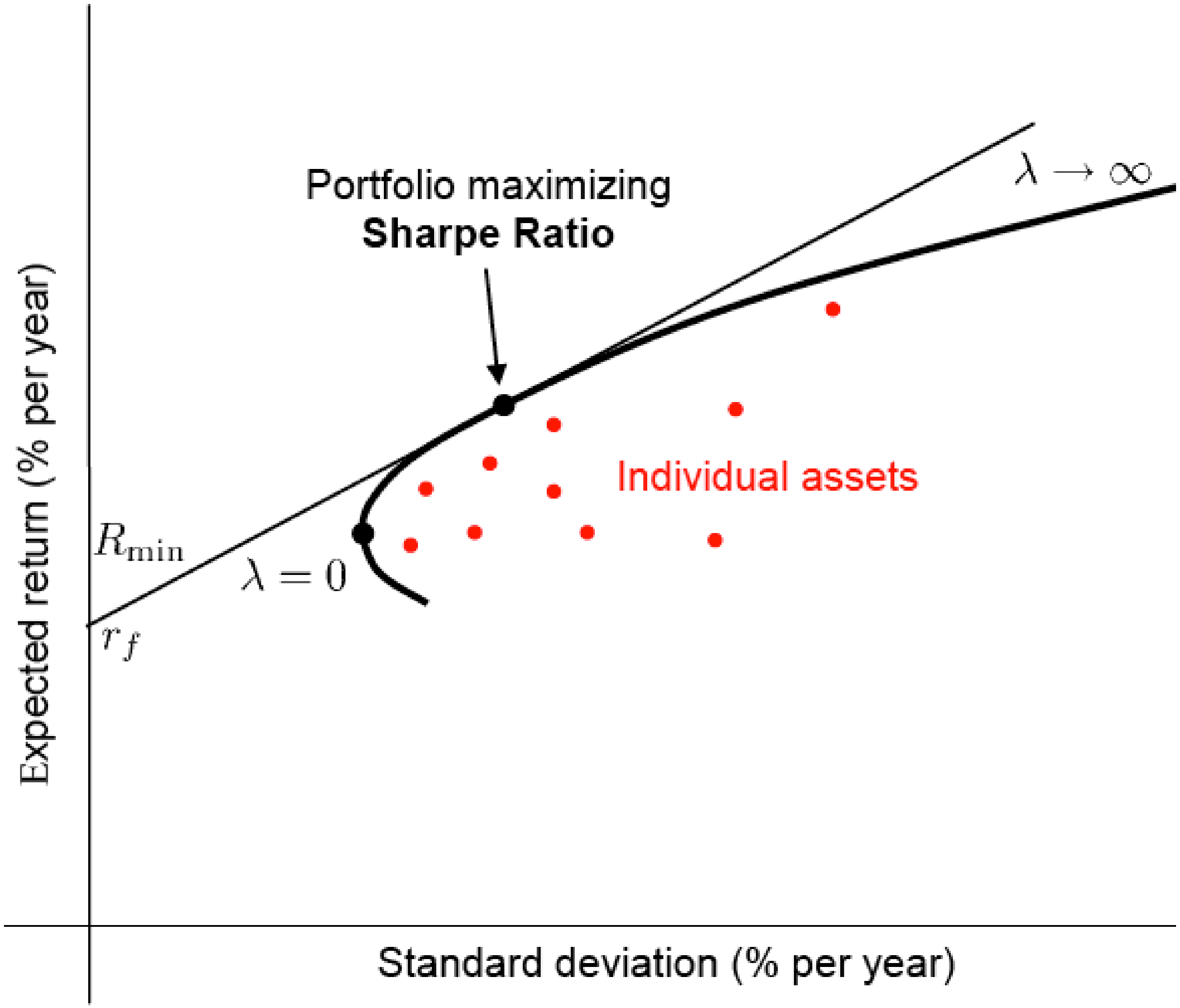
Mean-Variance Tradeoff

- The mean return $\mathbf{w}^T\boldsymbol{\mu}$ and the variance (risk) $\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$ constitute two important performance measures.

- Usually, the higher the mean return the higher the variance and vice-versa.

- Thus, we are faced with two objectives to be optimized. This is a multi-objective optimization problem.

- They define a fundamental mean-variance tradeoff curve (Pareto curve). The choice of a specific point in this tradeoff curve depends on how agressive or risk-averse the investor is.


Portfolio maximizing
Sharpe Ratio
$\lambda \to \infty$
Individual assets
$R_{\min}$
$\lambda = 0$
r_f
Expected return (% per year)
Standard deviation (% per year)

Sharpe Ratio

- The Sharpe ratio is akin to the SINR in communication systems.

- It is defined as the ratio of the mean return (excess w.r.t. the return of the risk-free asset r_f) to the risk measured as the square root of the variance (standard deviation):

$$\mathsf{S} = \frac{\mathbf{w}^T\boldsymbol{\mu} - r_f}{\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}}.$$

- The Sharpe ratio can be understood as the expected return per unit of risk.

- The portfolio $\mathbf{w}$ that maximizes the Sharpe ratio lies on the Pareto curve.

Mean-Variance Optimization (Markowitz, 1956)

- There are two obvious formulations for the portfolio optimization.
- Maximization of mean return:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\ \text{subject to} & \mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w} \leq \alpha \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

- Minimization of risk:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w} \\ \text{subject to} & \mathbf{w}^T\boldsymbol{\mu} \geq \beta \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

- Another obvious formulation is the scalarization of the multi-objective optimization problem:

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \gamma\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}\\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1.\end{array}$$

- These three formulations give different points on the Pareto optimal curve.

- They all require knowledge of one parameter (α, β, and γ).

- If we consider the Sharpe ratio as a good measure of performance, we could consider instead maximizing it directly.

- Additional constraints can be included such as $\mathbf{w} \geq \mathbf{0}$ to allow for long-only transactions (no short-selling).

Max Sharpe Ratio Portfolio Formulation

- Let's start by formulating the maximization of the Sharpe ratio as follows:

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{minimize}} & \sqrt{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}/\left(\mathbf{w}^T\boldsymbol{\mu}-r_f\right)\\ \text{subject to} & \mathbf{1}^T\mathbf{w}=1\\ & \mathbf{w}\geq 0\end{array}$$

- This is a quasi-convex problem as can be seen in the epigraph form:

$$\begin{array}{ll}\underset{\mathbf{w},t}{\text{minimize}} & t\\ \text{subject to} & t\left(\mathbf{w}^T\boldsymbol{\mu}-r_f\right)\geq\left\|\boldsymbol{\Sigma}^{1/2}\mathbf{w}\right\|\\ & \mathbf{1}^T\mathbf{w}=1\\ & \mathbf{w}\geq 0\end{array}$$

which can be solve by bisection over t.

Sharpe Ratio Formulation in Convex Form

- **Theorem**: The maximization of the Sharpe ratio can be rewritten in convex form as the QP

$$\begin{array}{ll} \underset{\tilde{\mathbf{w}}}{\text{minimize}} & \tilde{\mathbf{w}}^T \boldsymbol{\Sigma} \tilde{\mathbf{w}} \\ \text{subject to} & \left(\boldsymbol{\mu} - r_f \mathbf{1}\right)^T \tilde{\mathbf{w}} = 1 \\ & \mathbf{1}^T \tilde{\mathbf{w}} \geq 0 \\ & \tilde{\mathbf{w}} \geq 0. \end{array}$$

Proof:

- Defining $\tilde{\mathbf{w}} = t\mathbf{w}$ with $t = 1/\left(\mathbf{w}^T \boldsymbol{\mu} - r_f\right) > 0$, the objective becomes $\sqrt{\tilde{\mathbf{w}}^T \boldsymbol{\Sigma} \tilde{\mathbf{w}}}$, the sum constraint becomes $\mathbf{1}^T \tilde{\mathbf{w}} = t$, and the problem becomes

$$
\begin{array}{ll}
\underset{\mathbf{w},\tilde{\mathbf{w}},t}{\text{minimize}} & \sqrt{\tilde{\mathbf{w}}^T\boldsymbol{\Sigma}\tilde{\mathbf{w}}} \\
\text{subject to} & t = 1/\left(\mathbf{w}^T\boldsymbol{\mu} - r_f\right) > 0 \\
& \tilde{\mathbf{w}} = t\mathbf{w} \\
& \mathbf{1}^T\tilde{\mathbf{w}} = t > 0 \\
& \tilde{\mathbf{w}} \geq 0.
\end{array}
$$

- Observe that the first constraint $1/t = \mathbf{w}^T\boldsymbol{\mu} - r_f$ can be rewritten in terms of $\tilde{\mathbf{w}}$ as $1 = (\boldsymbol{\mu} - r_f\mathbf{1})^T\,\tilde{\mathbf{w}}$. So the problem becomes

$$
\begin{array}{ll}
\underset{\mathbf{w},\tilde{\mathbf{w}},t}{\text{minimize}} & \sqrt{\tilde{\mathbf{w}}^T\boldsymbol{\Sigma}\tilde{\mathbf{w}}} \\
\text{subject to} & (\boldsymbol{\mu} - r_f\mathbf{1})^T\,\tilde{\mathbf{w}} = 1 \\
& \tilde{\mathbf{w}} = t\mathbf{w} \\
& \mathbf{1}^T\tilde{\mathbf{w}} = t > 0 \\
& \tilde{\mathbf{w}} \geq 0.
\end{array}
$$

- Note that the strict inequality $t > 0$ is equivalent to $t \geq 0$ because $t = 0$ can never happen as $\tilde{\mathbf{w}}$ would be zero and the first constraint would not be satisfied.

- We can now get rid of $\mathbf{w}$ and t in the formulation as they can be directly obtained as $\mathbf{w} = \tilde{\mathbf{w}}/t$ and $t = \mathbf{1}^T\tilde{\mathbf{w}}$:

$$
\begin{array}{ll}
\underset{\tilde{\mathbf{w}}}{\text{minimize}} & \tilde{\mathbf{w}}^T \boldsymbol{\Sigma} \tilde{\mathbf{w}} \\
\text{subject to} & (\boldsymbol{\mu} - r_f \mathbf{1})^T \tilde{\mathbf{w}} = 1 \\
& \mathbf{1}^T \tilde{\mathbf{w}} \geq 0 \\
& \tilde{\mathbf{w}} \geq 0.
\end{array}
$$

- QED!

Value-at-Risk (VaR & CVaR) Models

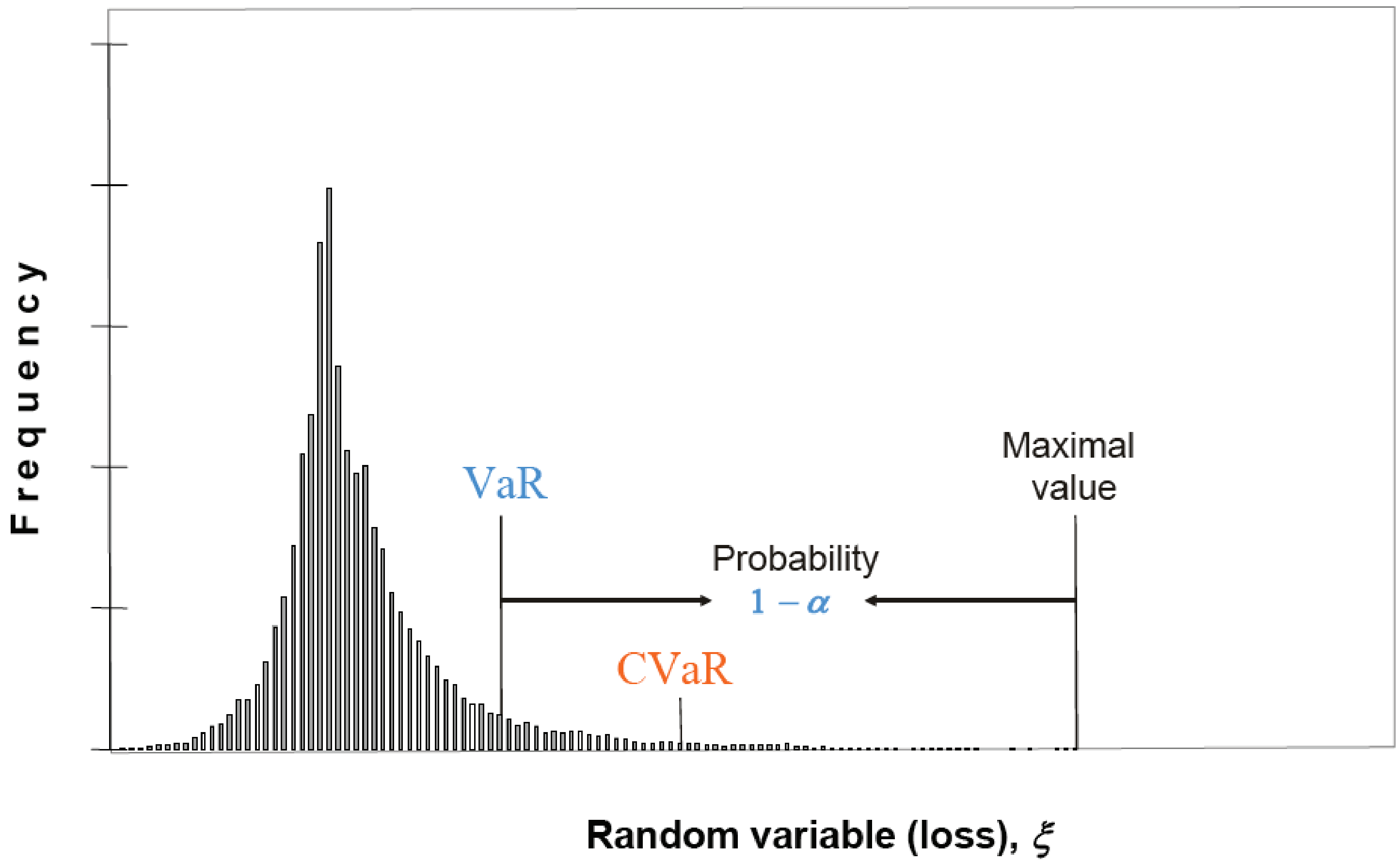
- Mean-variance model penalizes up-side and down-side risk equally, whereas most investors don't mind up-side risk.
- Solution: use alternative risk measures (not variance).
- VaR denotes the maximum loss with a specified confidence level (e.g., confidence level $= 95\%$, period $= 1$ day).
- Let ξ be a random variable representing the loss from a portfolio over some period of time:

$$\mathsf{VaR}_\alpha \;=\; \min\left\{\xi_0 : \mathsf{Pr}\left(\xi \leq \xi_0\right) \geq \alpha\right\}.$$

- Undesirable properties of VaR: does not take into account risks exceeding VaR, is nonconvex, is not sub-additive.

- The Conditional VaR (CVaR) takes into account the shape of the losses exceeding the VaR through the average:

$$\mathsf{CVaR}_{\alpha} \;=\; E\left[\xi \mid \xi \geq \mathsf{VaR}_{\alpha}\right].$$



CVaR Portfolio Formulation

- Dealing directly with VaR and CVaR quantities is not tractable.
- Let $f(\mathbf{w}, \mathbf{r})$ be an arbitrary cost function, where $\mathbf{w}$ is the optimization variable (portfolio) and $\mathbf{r}$ denotes the random parameters (asset returns). Example: $f(\mathbf{w}, \mathbf{r}) = -\mathbf{w}^T \boldsymbol{r}$.
- Consider, for example, the maximization of the mean return subject to a CVaR risk constraint on the loss:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T \boldsymbol{\mu} \\ \text{subject to} & \text{CVaR}_\alpha \left(f(\mathbf{w}, \mathbf{r})\right) \leq c \\ & \mathbf{1}^T \mathbf{w} = 1 \end{array}$$

where

$$\text{CVaR}_\alpha \left(f(\mathbf{w}, \mathbf{r})\right) = E\left[f(\mathbf{w}, \mathbf{r}) \mid f(\mathbf{w}, \mathbf{r}) \geq \text{VaR}_\alpha \left(f(\mathbf{w}, \mathbf{r})\right)\right].$$

CVaR in Convex Form

- **Theorem**: Define the auxiliary function

$$F_{\alpha}\left(\mathbf{w},\zeta\right)=\zeta+\frac{1}{1-\alpha}\int\left[f\left(\mathbf{w},\mathbf{r}\right)-\zeta\right]^{+}p\left(\mathbf{r}\right)d\mathbf{r}.$$

Then, we have the following two results:

(a) VaR_{α} is a minimizer of $F_{\alpha}\left(\mathbf{w},\zeta\right)$ with respect to ζ:

$$\mathsf{VaR}_{\alpha}\left(f\left(\mathbf{w},\mathbf{r}\right)\right)=\arg\min_{\zeta}F_{\alpha}\left(\mathbf{w},\zeta\right)$$

(b) CVaR_{α} equals minimal value (w.r.t. ζ) of $F_{\alpha}\left(\mathbf{w},\zeta\right)$:

$$\mathsf{CVaR}_{\alpha}\left(f\left(\mathbf{w},\mathbf{r}\right)\right)=\min_{\zeta}F_{\alpha}\left(\mathbf{w},\zeta\right)$$

Proof CVaR in Convex Form

(a) The minimizer $\zeta^\star$ of $F_\alpha(\mathbf{w},\zeta)$ satisfies: $0 \in \partial_\zeta F_\alpha(\mathbf{w},\zeta^\star)$. For example we choose the following subgradient:

$$\begin{aligned} s_\zeta F_\alpha(\mathbf{w},\zeta^\star) &= 1 - \frac{1}{1-\alpha}\int \mathcal{I}\left(f(\mathbf{w},\mathbf{r}) \geq \zeta^\star\right) p(\mathbf{r})\, d\mathbf{r} \\ &= 1 - \frac{1}{1-\alpha}\mathsf{Pr}\left(f(\mathbf{w},\mathbf{r}) \geq \zeta^\star\right) = 0 \end{aligned}$$

where $\mathcal{I}(\cdot)$ is the indicator function. Consequently,

$$\mathsf{Pr}\left(f(\mathbf{w},\mathbf{r}) \geq \zeta^\star\right) = 1-\alpha \Longrightarrow \zeta^\star = \mathsf{VaR}_\alpha\left(f(\mathbf{w},\mathbf{r})\right).$$

Proof CVaR in Convex Form (cont'd)

(b)

$$\min_{\zeta} F_{\alpha}\left(\mathbf{w},\zeta\right) = F_{\alpha}\left(\mathbf{w},\zeta^{\star}\right) = \zeta^{\star} + \frac{1}{1-\alpha}\int\left[f\left(\mathbf{w},\mathbf{r}\right)-\zeta^{\star}\right]^{+} p\left(\mathbf{r}\right)d\mathbf{r}.$$

Recall that

$$\begin{aligned}
\mathsf{CVaR}_{\alpha}\left(f\left(\mathbf{w},\mathbf{r}\right)\right) &= E\left[f\left(\mathbf{w},\mathbf{r}\right) \mid f\left(\mathbf{w},\mathbf{r}\right)\geq \mathsf{VaR}_{\alpha}\left(f\left(\mathbf{w},\mathbf{r}\right)\right)\right] \\
&= \frac{1}{1-\alpha}\int_{\mathbf{r}:f\left(\mathbf{w},\mathbf{r}\right)\geq \mathsf{VaR}_{\alpha}} f\left(\mathbf{w},\mathbf{r}\right)p\left(\mathbf{r}\right)d\mathbf{r} \\
&= \frac{1}{1-\alpha}\int\left[f\left(\mathbf{w},\mathbf{r}\right)-\mathsf{VaR}_{\alpha}\right]^{+} p\left(\mathbf{r}\right)d\mathbf{r} + \mathsf{VaR}_{\alpha}
\end{aligned}$$

CVaR in Convex Form (cont'd)

- **Corollary**:

$$\min_{\mathbf{w}} \mathsf{CVaR}_{\alpha}\left(f\left(\mathbf{w},\mathbf{r}\right)\right) = \min_{\mathbf{w},\zeta} F_{\alpha}\left(\mathbf{w},\zeta\right)$$

- In words, minimizing $F_{\alpha}\left(\mathbf{w},\zeta\right)$ simultaneously calculates the optimal CVaR and VaR.

- **Corollary**: If $f\left(\mathbf{w},\mathbf{r}\right)$ is convex in $\mathbf{w}$ for each $\mathbf{r}$, then $F_{\alpha}\left(\mathbf{w},\zeta\right)$ is convex!

Proof:

$$F_{\alpha}\left(\mathbf{w},\zeta\right) = \zeta + \frac{1}{1-\alpha}\int\left[f\left(\mathbf{w},\mathbf{r}\right)-\zeta\right]^{+} p\left(\mathbf{r}\right) d\mathbf{r}.$$

Reduction of CVaR Optimization to LP

- We start by considering discrete distributions or approximating the continuous one by a discrete one so that:

$$F_\alpha\left(\mathbf{w},\zeta\right)=\zeta+\frac{1}{1-\alpha}\sum_{k=1}^{N}p_k\left[f\left(\mathbf{w},\mathbf{r}^k\right)-\zeta\right]^+.$$

- Then include dummy variables z_k:

$$z_k\geq\left[f\left(\mathbf{w},\mathbf{r}^k\right)-\zeta\right]^+\Longrightarrow z_k\geq f\left(\mathbf{w},\mathbf{r}^k\right)-\zeta,\ z_k\geq 0$$

- The problem of minimizing the CVaR reduces to the following LP:

$$\begin{array}{ll} \underset{\mathbf{w},\zeta,\mathbf{z}}{\text{minimize}} & \zeta + \frac{1}{1-\alpha}\sum_{k=1}^{N} p_k z_k \\ \text{subject to} & f\left(\mathbf{w}, \mathbf{r}^k\right) - \zeta \leq z_k \geq 0 \\ & \mathbf{1}^T\mathbf{w} = 1 \end{array}$$

  where the cost function is assumed linear in $\mathbf{w}$: $f\left(\mathbf{w}, \mathbf{r}\right) = -\mathbf{w}^T\boldsymbol{r}$.

- The maximization of the mean return subject to a CVaR constraint becomes:

$$\begin{array}{ll} \underset{\mathbf{w},\zeta,\mathbf{z}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\ \text{subject to} & \zeta + \frac{1}{1-\alpha}\sum_{k=1}^{N} p_k z_k \leq c \\ & f\left(\mathbf{w}, \mathbf{r}^k\right) - \zeta \leq z_k \geq 0 \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

Robust Porfolio Design

- In practice, as usual, the parameters defining the optimization problem are not always perfectly known. Hence, the concept of robust design. We will consider worst-case robust design.

- There are many ways to formulate a worst-case design for the portfolio design. For example, we could consider a robust mean-variance formulation:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \underset{\boldsymbol{\Sigma} \in \mathcal{S}_{\boldsymbol{\Sigma}}}{\max} \; \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} \\ \text{subject to} & \underset{\boldsymbol{\mu} \in \mathcal{S}_{\boldsymbol{\mu}}}{\min} \; \mathbf{w}^T \boldsymbol{\mu} \geq \beta \\ & \mathbf{1}^T \mathbf{w} = 1. \end{array}$$

- Now, depending how we define the uncertainty sets for the mean return vector $\mathcal{S}_{\boldsymbol{\mu}}$ and for the covariance matrix $\mathcal{S}_{\boldsymbol{\Sigma}}$, the problem may be convex or not.

- We will consider one particular example based on modeling the returns via a factor model:

$$\mathbf{r} = \boldsymbol{\mu} + \mathbf{V}^T \mathbf{f} + \boldsymbol{\epsilon}$$

  where $\mathbf{f}$ denotes the random factors distributed according to $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, \mathbf{F})$ and $\boldsymbol{\epsilon}$ denotes the a random residual error with uncorrelated elements distributed according to $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$ with $\mathbf{D} = \mathrm{diag}(\mathbf{d})$.

- The returns are then distributed according to $\mathbf{r} \sim \mathcal{N}\left(\boldsymbol{\mu}, \mathbf{V}^T \mathbf{F} \mathbf{V} + \mathbf{D}\right)$ and the obtained return using portfolio $\mathbf{w}$ has mean $\boldsymbol{\mu}^T \mathbf{w}$ and variance $\mathbf{w}^T \left(\mathbf{V}^T \mathbf{F} \mathbf{V} + \mathbf{D}\right) \mathbf{w}$.

- We will consider uncertainty in the knowledge of $\boldsymbol{\mu}$, $\mathbf{V}$, and $\mathbf{D}$, while $\mathbf{F}$ is assumed know; in fact, we consider $\mathbf{F} = \mathbf{I}$.

- We can then formulate the robust mean-variance problem as

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \underset{\mathbf{V}\in\mathcal{S}_{\mathbf{V}},\mathbf{D}\in\mathcal{S}_{\mathbf{D}}}{\max} \ \mathbf{w}^T\left(\mathbf{V}^T\mathbf{V}+\mathbf{D}\right)\mathbf{w} \\ \text{subject to} & \underset{\boldsymbol{\mu}\in\mathcal{S}_{\boldsymbol{\mu}}}{\min}\ \mathbf{w}^T\boldsymbol{\mu}\geq\beta \\ & \mathbf{1}^T\mathbf{w}=1. \end{array}$$

- We will define the uncertainty as follows:

$$\begin{array}{rcl} \mathcal{S}_{\boldsymbol{\mu}} & = & \{\boldsymbol{\mu}\mid\boldsymbol{\mu}=\boldsymbol{\mu}_0+\boldsymbol{\delta},\,|\delta_i|\leq\gamma_i,\,i-1,\ldots,M\} \\ \mathcal{S}_{\mathbf{V}} & = & \{\mathbf{V}\mid\mathbf{V}=\mathbf{V}_0+\boldsymbol{\Delta},\,\|\boldsymbol{\Delta}\|_F\leq\rho\} \\ \mathcal{S}_{\mathbf{D}} & = & \left\{\mathbf{D}\mid\mathbf{D}=\text{diag}\left(\mathbf{d}\right),\,d_i\in\left[\underline{d}_i,\overline{d}_i\right],\,i-1,\ldots,M\right\}. \end{array}$$

- Let's now elaborate on each of the inner optimizations...

- First, consider the worst case mean return:

$$\min_{\boldsymbol{\mu} \in \mathcal{S}_{\boldsymbol{\mu}}} \boldsymbol{\mu}^T \mathbf{w} = \boldsymbol{\mu}_0^T \mathbf{w} + \min_{|\delta_i| \leq \gamma_i} \boldsymbol{\delta}^T \mathbf{w} = \boldsymbol{\mu}_0^T \mathbf{w} - \boldsymbol{\gamma}^T |\mathbf{w}|$$

which is a concave function.

- Second, let's turn to the second term of the worst-case variance:

$$\max_{\mathbf{D} \in \mathcal{S}_{\mathbf{D}}} \mathbf{w}^T \mathbf{D} \mathbf{w} = \max_{d_i \in \left[\underline{d}_i, \overline{d}_i\right]} \sum_{i=1}^{M} d_i w_i^2 = \sum_{i=1}^{M} \overline{d}_i w_i^2 = \mathbf{w}^T \overline{\mathbf{D}} \mathbf{w}$$

- Finally, let's focus on the first term of the worst-case variance:

$$\max_{\mathbf{V} \in \mathcal{S}_{\mathbf{V}}} \mathbf{w}^T \mathbf{V}^T \mathbf{V} \mathbf{w} \equiv \max_{\|\boldsymbol{\Delta}\|_F \leq \rho} \|\mathbf{V}_0 \mathbf{w} + \boldsymbol{\Delta} \mathbf{w}\| = \|\mathbf{V}_0 \mathbf{w}\| + \rho \|\mathbf{w}\|$$

- **Proof**: From the triangle inequality we have

$$\begin{aligned}\|\mathbf{V}_0\mathbf{w}+\boldsymbol{\Delta}\mathbf{w}\| &\leq \|\mathbf{V}_0\mathbf{w}\|+\|\boldsymbol{\Delta}\mathbf{w}\|\\ &\leq \|\mathbf{V}_0\mathbf{w}\|+\sqrt{\mathbf{w}^T\boldsymbol{\Delta}^T\boldsymbol{\Delta}\mathbf{w}}\\ &\leq \|\mathbf{V}_0\mathbf{w}\|+\|\mathbf{w}\|\,\|\boldsymbol{\Delta}\|_F\\ &\leq \|\mathbf{V}_0\mathbf{w}\|+\|\mathbf{w}\|\,\rho\end{aligned}$$

but this upper bound is achievable by the worst-case variable

$$\boldsymbol{\Delta}=\mathbf{u}\frac{\mathbf{w}^T}{\|\mathbf{w}\|}\rho$$

where

$$\mathbf{u}=\begin{cases}\frac{\mathbf{V}_0\mathbf{w}}{\|\mathbf{V}_0\mathbf{w}\|} & \text{if } \mathbf{V}_0\mathbf{w}\neq\mathbf{0}\\ \text{any unitary vector} & \text{otherwise.}\end{cases}$$

- Finally, the robust portfolio formulation is

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \left(\left\|\mathbf{V}_0\mathbf{w}\right\| + \rho\left\|\mathbf{w}\right\|\right)^2 + \mathbf{w}^T\overline{\mathbf{D}}\mathbf{w} \\ \text{subject to} & \boldsymbol{\mu}_0^T\mathbf{w} - \boldsymbol{\gamma}^T\left|\mathbf{w}\right| \geq \beta \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

or, better, as the SOCP

$$\begin{array}{ll} \underset{\mathbf{w},t}{\text{minimize}} & t^2 + \mathbf{w}^T\overline{\mathbf{D}}\mathbf{w} \\ \text{subject to} & t \geq \left\|\mathbf{V}_0\mathbf{w}\right\| + \rho\left\|\mathbf{w}\right\| \\ & \boldsymbol{\mu}_0^T\mathbf{w} \geq \beta + \boldsymbol{\gamma}^T\left|\mathbf{w}\right| \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

Index Replication

- **Index tracking** or **benchmark replication** is an strategy investment aimed at mimicking the risk/return profile of a financial instrument.

- For practical reasons, the strategy focuses on a **reduced basket** of representative securities.

- This problem is also regarded as **portfolio compression** and it is intimately related to compressed sensing and ℓ_1-norm minimization techniques.

- One example is the replication of an index, e.g., Hang Seng index, based on a reduced set of assets.

Tracking Error

- Let $\mathbf{c} \in \mathbb{R}^M$ represent the actual benchmark weight vector and let $\mathbf{w} \in \mathbb{R}^M$ denote the replicating portfolio.

- Investment managers seek to minimize the following **tracking error** performance measure:

$$TE_1(\mathbf{w}) = \sqrt{(\mathbf{c} - \mathbf{w})^T \boldsymbol{\Sigma} (\mathbf{c} - \mathbf{w})}.$$

- In practice, however, the benchmark weight vector $\mathbf{c}$ is unknown and the error measure is defined in terms of market observations.

Empirical Tracking Error

- Let $\mathbf{r}_c \in \mathbb{R}^N$ contain N temporal observations of the returns of the benchmark or index and let matrix $\mathbf{R} = \left[\begin{array}{ccc} \mathbf{r}_1 & \cdots & \mathbf{r}_N \end{array} \right] \in \mathbb{R}^{M \times N}$ contain column-wise the returns of the individual assets over time.

- The empirical tracking error can be defined as

$$TE_2(\mathbf{w}) = \left\| \mathbf{r}_c - \mathbf{R}^T \mathbf{w} \right\|_2.$$

- We can then formulate the **sparse index replication** problem as

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & TE_2(\mathbf{w}) + \gamma \text{card}(\mathbf{w}) \\ \text{subject to} & \mathbf{w} \geq \mathbf{0} \\ & \mathbf{1}^T \mathbf{w} = 1. \end{array}$$

Index Replication with ℓ_1-Norm Minimization

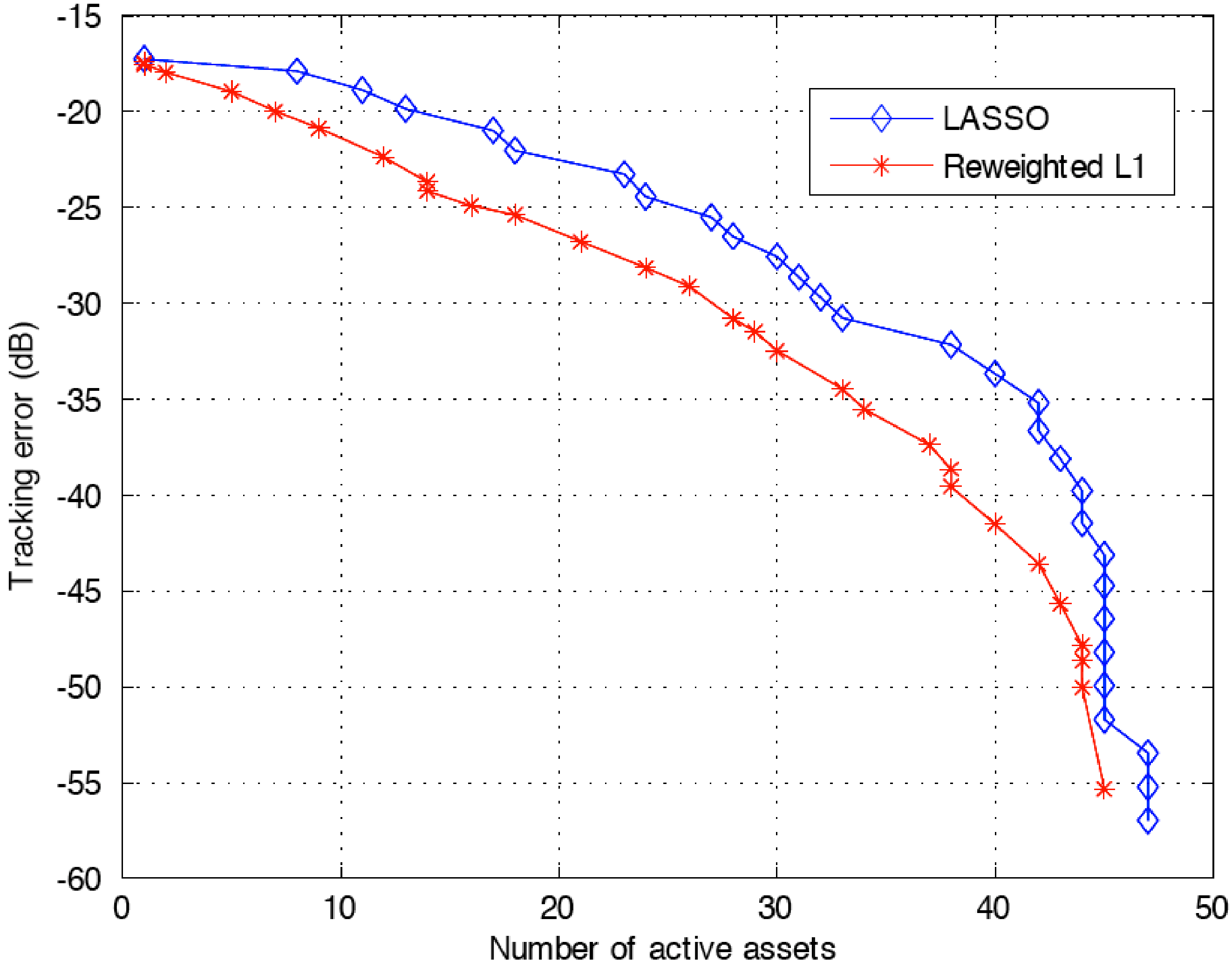
- The cardinality operator is nonconvex and can be approximated in a number of ways.

- The simplest approximation is based on the ℓ_1-norm:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \left\|\mathbf{r}_c - \mathbf{R}^T\mathbf{w}\right\|_2 + \gamma \left\|\mathbf{w}\right\|_1 \\ \text{subject to} & \mathbf{w} \geq \mathbf{0} \\ & \mathbf{1}^T\mathbf{w} = 1. \end{array}$$

- Of course, there are more sophisticated methods such as the reweighted ℓ_1-norm approximation based on a successive convex approximation (see lecture on ℓ_1-norm minimization for details).

Simulations of Sparse Index Replication



Summary

- We have introduced basic concepts and data model for portfolio optimization.
- Then, we have considered different formulations for the portfolio optimization problem:
  - basic mean-variance formulations
  - Sharpe ratio maximization in convex form
  - CVaR optimization in convex form
  - worst-case robust designs in convex form
- Finally, we have studied related problem such as the index replication based on the ℓ_1-norm.

References

- Yiyong Feng and Daniel P. Palomar, *A Signal Processing Perspective on Financial Engineering*, Foundations and Trends® in Signal Processing, Now Publishers, vol. 9, no. 1-2, 2016.

- D. Luenberger, *Investment Science*, Oxford Univ. Press, 1997.

- D. Ruppert, *Statistics and Data Analysis for Financial Engineering*, Springer Texts in Statistics, 2010.

- G. Cornuejols and R. Tutuncu, *Optimization Methods in Finance*, Cambridge Univ. Press, 2007.

- F. Fabozzi, P. Kolm, D. Pachamanova, S. Focardi, *Robust Portfolio Optimization and Management*, Wiley Finance, 2007.