Portfolio Optimization with Alternative Risk Measures

Prof. Daniel P. Palomar
The Hong Kong University of Science and Technology (HKUST)

MAFS6010R- Portfolio Optimization with R
MSc in Financial Mathematics
Fall 2018-19, HKUST, Hong Kong

Outline

Outline

Motivation

- The Markowitz portfolio has never been embraced by practitioners, among other reasons because
  1. variance is not a good measure of risk in practice since it penalizes both the unwanted high losses and the desired low losses: the solution is to use **alternative measures for risk, e.g., VaR and CVaR**,
  2. it is highly sensitive to parameter estimation errors (i.e., to the covariance matrix $\boldsymbol{\Sigma}$ and especially to the mean vector $\boldsymbol{\mu}$): solution is **robust optimization**,
  3. it only considers the risk of the portfolio as a whole and ignores the risk diversification: solution is the **risk-parity portfolio**.

👉 ***We will here consider more meaningful measures for risk than the variance, like the downside risk (DR), Value-at-Risk (VaR), Conditional VaR (CVaR) or Expected Shortfall (ES), and drawdown (DD).***

Outline

Outline

Returns

- Let us denote the log-returns of N assets at time t with the vector $\mathbf{r}_t \in \mathbb{R}^N$.
- The time index t can denote any arbitrary period such as days, weeks, months, 5-min intervals, etc.
- $\mathcal{F}_{t-1}$ denotes the previous historical data.
- Econometrics aims at modeling $\mathbf{r}_t$ conditional on $\mathcal{F}_{t-1}$.
- $\mathbf{r}_t$ is a multivariate stochastic process with conditional mean and covariance matrix denoted as[1]

$$\begin{aligned}\boldsymbol{\mu}_t &\triangleq \mathsf{E}\left[\mathbf{r}_t \mid \mathcal{F}_{t-1}\right]\\ \boldsymbol{\Sigma}_t &\triangleq \mathrm{Cov}\left[\mathbf{r}_t \mid \mathcal{F}_{t-1}\right] = \mathsf{E}\left[(\mathbf{r}_t - \boldsymbol{\mu}_t)(\mathbf{r}_t - \boldsymbol{\mu}_t)^T \mid \mathcal{F}_{t-1}\right].\end{aligned}$$

[1] Y. Feng and D. P. Palomar, *A Signal Processing Perspective on Financial Engineering*. Foundations and Trends in Signal Processing, Now Publishers, 2016.

I.I.D. Model

- For simplicity we will assume that $\mathbf{r}_t$ follows an i.i.d. distribution (which is not very innacurate in general).
- That is, both the conditional mean and conditional covariance are constant

$$\boldsymbol{\mu}_t = \boldsymbol{\mu},$$
$$\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}.$$

- Very simple model, however, it is one of the most fundamental assumptions for many important works, e.g., the Nobel prize-winning Markowitz portfolio theory[2].

[2] H. Markowitz, "Portfolio selection," *J. Financ.*, vol. 7, no. 1, pp. 77–91, 1952.

Parameter Estimation

- Consider the i.i.d. model:

$$\mathbf{r}_t = \boldsymbol{\mu} + \mathbf{w}_t,$$

  where $\boldsymbol{\mu} \in \mathbb{R}^N$ is the mean and $\mathbf{w}_t \in \mathbb{R}^N$ is an i.i.d. process with zero mean and constant covariance matrix $\boldsymbol{\Sigma}$.
- The mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ have to be estimated in practice based on T observations.
- The simplest estimator is the sample estimator:
  - sample mean estimator: $\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t$
  - sample covariance matrix: $\hat{\boldsymbol{\Sigma}} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{r}_t - \hat{\boldsymbol{\mu}})(\mathbf{r}_t - \hat{\boldsymbol{\mu}})^T$.
- Many more sophisticated estimators exist, namely: shrinkage estimators, Black-Litterman estimators, etc.

Outline

Portfolio Return

- Suppose the budget is B dollars.
- The portfolio $\mathbf{w} \in \mathbb{R}^N$ denotes the normalized weights of the assets such that $\mathbf{1}^T\mathbf{w} = 1$ (then $B\mathbf{w}$ denotes dollars invested in the assets).
- For each asset, the initial wealth is Bw_i and the end wealth is

$$Bw_i\left(p_{i,t}/p_{i,t-1}\right) = Bw_i\left(R_{it} + 1\right).$$

- Then the portfolio return is

$$R_t^p = \frac{\sum_{i=1}^N Bw_i\left(R_{it} + 1\right) - B}{B} = \sum_{i=1}^N w_i R_{it} \approx \sum_{i=1}^N w_i r_{it} = \mathbf{w}^T\mathbf{r}_t$$

- The portfolio expected return and variance are $\mathbf{w}^T\boldsymbol{\mu}$ and $\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$, respectively.[3]

[3] G. Cornuejols and R. Tütüncü, *Optimization Methods in Finance*. Cambridge University Press, 2006.

Performance Measures

- Expected return: $\mathbf{w}^T\boldsymbol{\mu}$
- Volatility: $\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}$
- Sharpe Ratio (SR): expected return per unit of risk

$$\mathrm{SR} = \frac{\mathbf{w}^T\boldsymbol{\mu} - r_f}{\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}}$$

  where r_f is the risk-free rate (e.g., interest rate on a three-month U.S. Treasury bill).
- Information Ratio (IR):

$$\mathrm{IR} = \frac{\mathbf{w}^T\boldsymbol{\mu}}{\sqrt{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}}$$

- Drawdown (DD): decline from a historical peak of the cumulative profit $X(t)$: $D(t) = \max_{1\leq\tau\leq t} X(\tau) - X(t)$ (unnormalized)
- VaR (Value at Risk)
- ES (Expected Shortfall) or CVaR (Conditional Value at Risk)

Practical Constraints

- Capital budget constraint:

$$\mathbf{w}^T\mathbf{1} = 1.$$

- Long-only constraint:

$$\mathbf{w} \geq 0.$$

- Market-neutral constraint:

$$\mathbf{w}^T\mathbf{1} = 0.$$

- Turnover constraint:

$$\|\mathbf{w} - \mathbf{w}_0\|_1 \leq u$$

where $\mathbf{w}_0$ is the currently held portfolio.

Practical Constraints

- Holding constraint:

$$\mathbf{l} \leq \mathbf{w} \leq \mathbf{u}$$

  where $\mathbf{l} \in \mathbb{R}^N$ and $\mathbf{u} \in \mathbb{R}^N$ are lower and upper bounds of the asset positions, respectively.

- Cardinality constraint:

$$\|\mathbf{w}\|_0 \leq K.$$

- Leverage constraint:

$$\|\mathbf{w}\|_1 \leq 2.$$

Risk Control

- In finance, the expected return $\mathbf{w}^T\boldsymbol{\mu}$ is very relevant as it quantifies the average benefit.
- However, in practice, the average performance is not enough to characterize an investment 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[4]: 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 downside risk, VaR, ES, drawdown, etc.

[4] H. Markowitz, "Portfolio selection," *J. Financ.*, vol. 7, no. 1, pp. 77–91, 1952.

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: it 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.

Markowitz mean-variance portfolio (1952)

- The idea of the Markowitz framework[5] is to find a trade-off between the expected return $\mathbf{w}^T\boldsymbol{\mu}$ and the risk of the portfolio measured by the variance $\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$:

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

  where $\mathbf{w}^T\mathbf{1} = 1$ is the capital budget constraint and λ is a parameter that controls how risk-averse the investor is.

- This is a convex QP with only one linear constraint which admits a closed-form solution:

$$\mathbf{w}^\star = \frac{1}{2\lambda}\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\mu} + \nu^\star\mathbf{1}\right),$$

  where $\nu^\star$ is the optimal dual variable $\nu^\star = \frac{2\lambda - \mathbf{1}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{\mathbf{1}^T\boldsymbol{\Sigma}^{-1}\mathbf{1}}$.

[5] H. Markowitz, "Portfolio selection," *J. Financ.*, vol. 7, no. 1, pp. 77–91, 1952.

Global Minimum Variance Portfolio (GMVP)

- The global minimum variance portfolio (GMVP) ignores the expected return and focuses on the risk only:

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

- It is a simple convex QP with solution

$$\mathbf{w}_{\text{GMVP}} = \frac{1}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}} \boldsymbol{\Sigma}^{-1} \mathbf{1}.$$

- It is widely used in academic papers for simplicity of evaluation and comparison of different estimators of the covariance matrix $\boldsymbol{\Sigma}$ (while ignoring the estimation of $\boldsymbol{\mu}$).

Outline

Drawbacks of Markowitz's formulation

- The Markowitz portfolio has never been embraced by practitioners, among other reasons because
  1. variance is not a good measure of risk in practice since it penalizes both the unwanted high losses and the desired low losses: the solution is to use **alternative measures for risk, e.g., VaR and CVaR**,
  2. it is highly sensitive to parameter estimation errors (i.e., to the covariance matrix $\boldsymbol{\Sigma}$ and especially to the mean vector $\boldsymbol{\mu}$): solution is **robust optimization**,
  3. it only considers the risk of the portfolio as a whole and ignores the risk diversification: solution is the **risk-parity portfolio**.

👉 ***We will here consider more meaningful measures for risk than the variance, like the downside risk (DR), Value-at-Risk (VaR), Conditional VaR (CVaR) or Expected Shortfall (ES), and drawdown (DD).***

Outline

Variance as risk measure

- In finance, the mean return is very relevant as it quantifies the average benefit of the investment.
- 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 the variance as considered by Markowitz in 1952:[6] a higher variance means that there are large peaks in the risk distribution which may cause a big loss.
- However, Markowitz himself already recognized and stressed the limitations of the mean-variance analysis.[7]

[6] H. Markowitz, "Portfolio selection," *J. Financ.*, vol. 7, no. 1, pp. 77–91, 1952.

[7] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. Wiley, 1959.

Alternatives to variance as risk measure

- Variance is not a good measure of risk in practice since it penalizes both the unwanted high losses and the desired low losses (or gains).[8]
- Indeed, the mean-variance portfolio framework penalizes up-side and down-side risk equally, whereas most investors don't mind up-side risk.
- To overcome the limitations of the variance as risk measure, a number of alternative risk measures have been proposed, for example:
  - Downside Risk (DR)
  - Value-at-Risk (VaR)
  - Conditional Value-at-Risk (CVaR)
  - Drawdown (DD):
    - maximum DD
    - average DD
    - Conditional Drawdown at Risk (CDaR)

[8] A. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2005.

Downside Risk (DR)

- Let R be a random variable representing the return of an asset or portfolio (e.g., $R = \mathbf{w}^T\mathbf{r}$ where $\mathbf{r}$ denotes the vector of random returns of the assets).
- We are familiar with the mean return $\mu = \mathsf{E}[R]$ and with the variance $\sigma^2 = \mathsf{E}[(R-\mu)^2]$.
- The idea of downside risk is that the left-handside of the return distribution involves risk while the right-handside contains the better investment opportunities.
- Interest in downside risk arose in the early 1950s.
- One example is the semi-variance, already considered by Markowitz in 1959.[9]
- The semi-variance measures the variability of the returns below the mean.

[9] H. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*. Wiley, 1959.

LPM and semivariance

- The semivariance is a special case of the more general lower partial moments (LPM):
$$\mathrm{LPM} = \mathrm{E}\left[\left((\tau - R)^{+}\right)^{\alpha}\right],$$
where $(\cdot)^{+} = \max(0, \cdot)$.
- The parameter τ is termed the disaster level.
- The parameter α reflects the investor's feeling about the relative consequences of falling short of τ by various amounts:
    - the value $\alpha = 1$ (which suits a neutral investor) separates risk-seeking ($0 < \alpha < 1$) from risk-averse ($\alpha > 1$) behavior with regard to returns below the target τ.
- By changing the parameters α and τ most downside measures used in practice can be formed.
- In particular, setting $\alpha = 2$ and $\tau = \mathrm{E}[R]$ yields the semi-variance (or lower partial variance):
$$\mathrm{SV} = \mathrm{E}\left[\left((E[R] - R)^{+}\right)^{2}\right].$$

Value-at-Risk (VaR)

- To overcome the drawback of variance, another popular single side risk measurement is the Value-at-Risk (VaR) initially proposed by J.P. Morgan.
- 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 (e.g., $\xi = -\mathbf{w}^T\mathbf{r}$ where $\mathbf{r}$ denotes the vector of random returns of the assets).
- The VaR is defined as

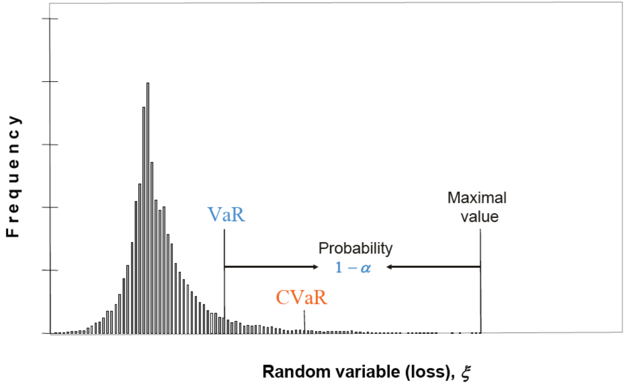
$$\text{VaR}_\alpha = \inf\left\{\xi_0 : \Pr\left(\xi \leq \xi_0\right) \geq \alpha\right\}$$

  with α the confidence level, say, $\alpha = 0.95$.
- However, this measure does not take into account losses exceeding VaR, is nonconvex, and is not subadditive.

Conditional Value-at-Risk (CVaR)

- The Conditional Value-at-Risk (CVaR) is also called Expected Shortfall (ES).
- The CVaR takes into account the shape of the losses exceeding the VaR through the average:

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



Drawdown

- The drawdown (DD) at time t is defined as the decline from a historical peak of the cumulative profit $X(t)$.
- The unnormalized version is

$$D(t)^{\text{unnorm}} = \max_{1 \le \tau \le t} X(\tau) - X(t).$$

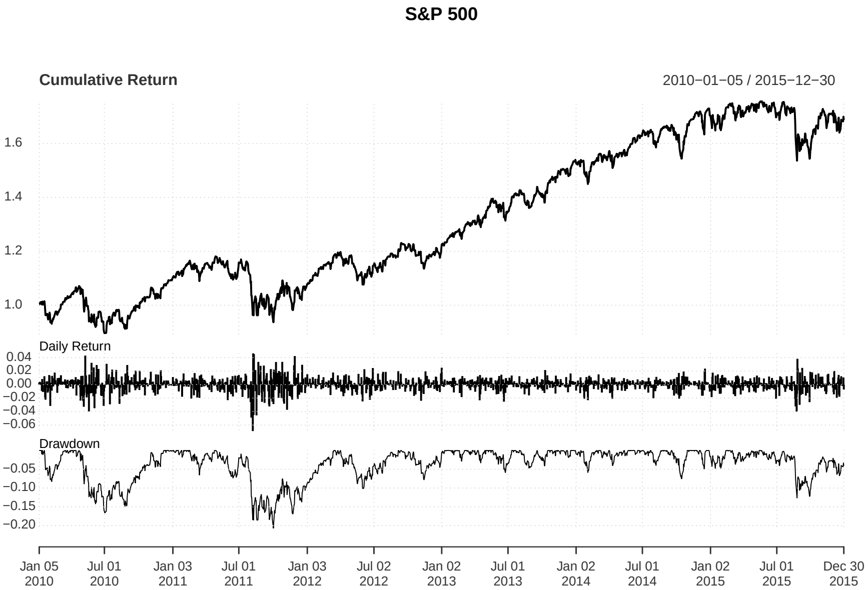
- But in practice, the normalized version is used:

$$D(t) = \frac{\text{HWM}_t - X(t)}{\text{HWM}_t}$$

where HWM_t is the high water mark of $X(t)$ defined as

$$\text{HWM}_t = \max_{1 \le \tau \le t} X(\tau).$$

Drawdown


S&P 500
Cumulative Return
2010−01−05 / 2015−12−30
1.6
1.4
1.2
1.0
Daily Return
0.04
0.02
0.00
−0.02
−0.04
−0.06
Drawdown
−0.05
−0.10
−0.15
−0.20
Jan 05 2010
Jul 01 2010
Jan 03 2011
Jul 01 2011
Jan 03 2012
Jul 02 2012
Jan 02 2013
Jul 01 2013
Jan 02 2014
Jul 01 2014
Jan 02 2015
Jul 01 2015
Dec 30 2015

Drawdown

- Then one can define the maximum DD (Max-DD) over a period $t = 1, \dots, T$ as
$$M(T) = \max_{1 \le t \le T} D(t)$$
- Also the average DD (Ave-DD) over a period $t = 1, \dots, T$ as
$$A(T) = \frac{1}{T} \sum_{1 \le t \le T} D(t)$$
- Similarly to the CVaR, we can define the Conditional Drawdown at Risk (CDaR) as the mean of the worst $100(1-\alpha)\%$ drawdowns.

Outline

Mean-downside risk portfolio

- Recall Markowitz mean-variance portfolio formulation:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda \mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w} \\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}. \end{array}$$

- Instead of using the variance we can use a downside risk measure, obtaining the mean-downside risk formulation (introduced in 1977).
- For example, the LPM can be approximated as

$$E\left[\left((\tau - R)^+\right)^\alpha\right] \approx \frac{1}{T}\sum_{t=1}^{T}\left((\tau - R_t)^+\right)^\alpha$$

where $R_t = \mathbf{w}^T\mathbf{r}_t$.
- The mean-LPM portfolio formulation is the convex (depending on α) problem

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda \frac{1}{T}\sum_{t=1}^{T}\left(\left(\tau - \mathbf{w}^T\mathbf{r}_t\right)^+\right)^\alpha \\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}. \end{array}$$

Mean-semivariance portfolio

- In particular, we can approximate the semivariance as

$$E\left[\left((E[R]-R)^+\right)^2\right] \approx \frac{1}{T}\sum_{t=1}^{T}\left((E[R]-R_t)^+\right)^2$$
$$\approx \frac{1}{T}\sum_{t=1}^{T}\left(\left(\frac{1}{T}\sum_{t=1}^{T}R_t - R_t\right)^+\right)^2$$

- The mean-semivariance portfolio formulation is the convex QP problem

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda\frac{1}{T}\sum_{t=1}^{T}\left(\left(\mathbf{w}^T\boldsymbol{\mu} - \mathbf{w}^T\mathbf{r}_t\right)^+\right)^2\\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

Mean-LPM portfolio portfolio with different α's

- For less risk-averse investors, we can consider the mean-LPM portfolio formulation with $\alpha = 1$, which is an LP:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{w}^T\boldsymbol{\mu} - \mathbf{w}^T\mathbf{r}_t\right)^+ \\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}. \end{array}$$

- For more risk-averse investors, we can consider the mean-LPM convex portfolio formulation with $\alpha = 3$:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda\frac{1}{T}\sum_{t=1}^{T}\left(\left(\mathbf{w}^T\boldsymbol{\mu} - \mathbf{w}^T\mathbf{r}_t\right)^+\right)^3 \\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}. \end{array}$$

Outline

Mean-CVaR portfolio

- A 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 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:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\ \text{subject to} & \text{CVaR}_\alpha\left(f(\mathbf{w},\mathbf{r})\right) \leq c \\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} \end{array}$$

where

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

CVaR portfolio

- Rockafellar and Uryasev[10] first proposed to minimize the CVaR of the portfolio loss as follows:

$$\begin{array}{ll} \underset{\mathbf{w}}{\text{minimize}} & \text{CVaR}_\alpha\left(\mathbf{w}^T\mathbf{r}\right) \\ \text{subject to} & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0} \end{array}$$

where

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

[10] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *J. Risk*, vol. 2, pp. 21–42, 2000.

CVaR in Convex Form

- Define the auxiliary convex function

$$F_\alpha(\mathbf{w},\zeta) = \zeta + \frac{1}{1-\alpha}\mathsf{E}[-\mathbf{w}^T\mathbf{r} - \zeta]^+,$$

  where $[x]^+ = \max(x, 0)$.
- Rockafellar and Uryasev show that

1. $\mathrm{VaR}_\alpha(-\mathbf{w}^T\mathbf{r})$ is a minimizer of $F_\alpha(\mathbf{w},\zeta)$ w.r.t. ζ:

$$\mathrm{VaR}_\alpha(-\mathbf{w}^T\mathbf{r}) \in \arg\min_\zeta F_\alpha(\mathbf{w},\zeta).$$

2. $\mathrm{CVaR}_\alpha(-\mathbf{w}^T\mathbf{r})$ equals minimum $F_\alpha(\mathbf{w},\zeta)$ w.r.t. ζ:

$$\mathrm{CVaR}_\alpha(-\mathbf{w}^T\mathbf{r}) = \min_\zeta F_\alpha(\mathbf{w},\zeta).$$

Proof CVaR in Convex Form

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

$$0 = s_\zeta F_\alpha(\mathbf{w}, \zeta^\star) = 1 - \frac{1}{1-\alpha}\int \mathbf{1}_{\{-\mathbf{w}^T\mathbf{r} > \zeta^\star\}} p(\mathbf{r}) d\mathbf{r}$$
$$= 1 - \frac{1}{1-\alpha} P\left(-\mathbf{w}^T\mathbf{r} > \zeta^\star\right),$$

where $\mathbf{1}$. is the indicator function. Solving the above equation, we have

$$P\left(-\mathbf{w}^T\mathbf{r} > \zeta^\star\right) = 1 - \alpha \Longrightarrow \zeta^\star = \mathrm{VaR}_\alpha(-\mathbf{w}^T\mathbf{r}).$$

Proof CVaR in Convex Form

2. First, we have

$$\min_{\zeta} F_{\alpha}(\mathbf{w}, \zeta) = F_{\alpha}(\mathbf{w}, \zeta^{\star}) = \zeta^{\star} + \frac{1}{1-\alpha}\mathsf{E}[-\mathbf{w}^T\mathbf{r} - \zeta^{\star}]^{+}.$$

Recall that

$$\begin{aligned}
\mathrm{CVaR}_{\alpha}(-\mathbf{w}^T\mathbf{r}) &= \mathsf{E}\left[-\mathbf{w}^T\mathbf{r} \middle| -\mathbf{w}^T\mathbf{r} > \mathrm{VaR}_{\alpha}(-\mathbf{w}^T\mathbf{r})\right] \\
&= \frac{1}{1-\alpha}\int_{-\mathbf{w}^T\mathbf{r} > \mathrm{VaR}_{\alpha}(-\mathbf{w}^T\mathbf{r})} \left(-\mathbf{w}^T\mathbf{r}\right) p(\mathbf{r}) d\mathbf{r} \\
&= \frac{1}{1-\alpha}\int \left[-\mathbf{w}^T\mathbf{r} - \mathrm{VaR}_{\alpha}(-\mathbf{w}^T\mathbf{r})\right]^{+} p(\mathbf{r}) d\mathbf{r} \\
&\quad + \mathrm{VaR}_{\alpha}(-\mathbf{w}^T\mathbf{r}).
\end{aligned}$$

CVaR in Convex Form

- Corollary:

$$\min_{\mathbf{w}} \mathrm{CVaR}_{\alpha}\left(-\mathbf{w}^{T}\mathbf{r}\right) = \min_{\mathbf{w},\zeta} F_{\alpha}\left(\mathbf{w},\zeta\right)$$

- In words, minimizing $F_{\alpha}\left(\mathbf{w},\zeta\right)$ simultaneously calculates the optimal CVaR and VaR.
- Corollary: Because $-\mathbf{w}^{T}\mathbf{r}$ is convex in $\mathbf{w}$ for each $\mathbf{r}$, then $F_{\alpha}\left(\mathbf{w},\zeta\right)$ is convex!

Proof:

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

Sample Average Approximation of CVaR

- Sample average approximation of $F_\alpha(\mathbf{w},\zeta)$:

$$F_\alpha(\mathbf{w},\zeta) = \zeta + \frac{1}{1-\alpha}\mathsf{E}[-\mathbf{w}^T\mathbf{r} - \zeta]^+$$
$$\approx \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T}[-\mathbf{w}^T\mathbf{r}_t - \zeta]^+.$$

CVaR portfolio as an LP

- We first include the dummy variables z_t:

$$z_t \geq \left[-\mathbf{w}^T\mathbf{r}_t - \zeta\right]^+ \Longrightarrow z_t \geq -\mathbf{w}^T\mathbf{r}_t - \zeta,\ z_t \geq 0$$

- CVaR portfolio problem can be approximated by an LP:

$$\begin{array}{ll}\underset{\mathbf{w},\{z_t\},\zeta}{\text{minimize}} & \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T} z_t \\ \text{subject to} & 0 \leq z_t \geq -\mathbf{w}^T\mathbf{r}_t - \zeta, \quad t = 1,\ldots,T \\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

Mean-CVaR portfolio as an LP

- We can also consider the maximization of the mean return subject to a CVaR constraint:

$$
\begin{array}{ll}
\underset{\mathbf{w},\{z_t\},\zeta}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\
\text{subject to} & \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T} z_t \leq c \\
 & 0 \leq z_t \geq -\mathbf{w}^T\mathbf{r}_t - \zeta, \quad t = 1, \ldots, T \\
 & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.
\end{array}
$$

- Or a mean-CVaR objective:

$$
\begin{array}{ll}
\underset{\mathbf{w},\{z_t\},\zeta}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} - \lambda\left(\zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T} z_t\right) \\
\text{subject to} & 0 \leq z_t \geq -\mathbf{w}^T\mathbf{r}_t - \zeta, \quad t = 1, \ldots, T \\
 & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.
\end{array}
$$

Outline

Drawdown (DD)

- Let $\mathbf{r}(t)$ be the return vector of the N stocks at time t.
- Define the cumulative (uncompounded) return vector as

$$\mathbf{r}^{\text{cum}}(t) = \sum_{\tau=1}^{t} \mathbf{r}(\tau)$$

  (Note: the compounded return is $\prod_{\tau=1}^{t}(\mathbf{1} + \mathbf{r}(\tau)) - \mathbf{1}$.)
- The portfolio return is $r_p(t) = \mathbf{w}^T \mathbf{r}(t)$ and cumulative return

$$r_p^{\text{cum}}(t) = \mathbf{w}^T \mathbf{r}^{\text{cum}}(t)$$

- The drawdown (DD) at time t can be written as

$$D(t) = \max_{1 \leq \tau \leq t} r_p^{\text{cum}}(\tau) - r_p^{\text{cum}}(t)$$

Max-DD, Ave-DD, and CDaR

- The maximum DD (Max-DD) over a period $t = 1, \dots, T$ is

$$M(T) = \max_{1 \leq t \leq T} D(t)$$

- The average DD (Ave-DD) over a period $t = 1, \dots, T$ is

$$A(T) = \frac{1}{T} \sum_{1 \leq t \leq T} D(t)$$

- Similarly to the CVaR, we can define the Conditional Drawdown at Risk (CDaR) as the mean of the worst $100(1-\alpha)\%$ drawdowns:

$$\Delta_{\alpha}(T) = \frac{1}{(1-\alpha)T} \sum_{t \in \Omega_{\alpha}} D(t),$$

where $\Omega_{\alpha} = \{1 \leq t \leq T \mid D(t) \geq \xi_{\alpha}\}$ with ξ_{α} being the threshold such that exactly $100(1-\alpha)\%$ of drawdowns exceeds that limit.

CDaR in Convex Form

- The CDaR can be conveniently expressed as[11]

$$\Delta_\alpha(\mathbf{w}) = \min_{\zeta}\{\zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T}[D_t(\mathbf{w}) - \zeta]^+\}$$

- When α tends to 1, the CDaR tends to the maximum drawdown, i.e.,

$$\Delta_1(T) = M(T)$$

- When α tends to 0, the CDaR tends to the average drawdown, i.e.,

$$\Delta_0(T) = A(T)$$

[11] A. Chekhlov, S. Uryasev, and M. Zabarankin, "Portfolio optimization with drawdown constraints," *Research Report 2000-5. Available at SSRN: https://ssrn.com/abstract=223323 or http://dx.doi.org/10.2139/ssrn.223323*, 2000.

Mean-Max-DD portfolio as an LP

- We can consider the maximization of the mean return subject to a Max-DD constraint:

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \max_{1\leq t\leq T}\{\max_{1\leq\tau\leq t}\mathbf{w}^T\mathbf{r}_\tau^{\text{cum}} - \mathbf{w}^T\mathbf{r}_t^{\text{cum}}\} \leq c\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w}\geq\mathbf{0}.\end{array}$$

- Removing one maximum operator is trivial:

$$\begin{array}{ll}\underset{\mathbf{w}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \max_{1\leq\tau\leq t}\mathbf{w}^T\mathbf{r}_\tau^{\text{cum}} - \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq c, \quad \forall 1\leq t\leq T\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w}\geq\mathbf{0}.\end{array}$$

Mean-Max-DD portfolio as an LP

- To remove the other max operator, we need to introduce some additional variables:

$$\begin{array}{ll}\underset{\mathbf{w},\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & u_t - \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq c, \quad \forall 1 \leq t \leq T\\ & u_t \geq \mathbf{w}^T\mathbf{r}_\tau^{\text{cum}} \qquad \forall 1 \leq t \leq T, 1 \leq \tau \leq t\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

- We can reduce the large number of constraints by rewriting it as

$$\begin{array}{ll}\underset{\mathbf{w},\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & u_t - \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq c, \quad \forall 1 \leq t \leq T\\ & u_t \geq \mathbf{w}^T\mathbf{r}_t^{\text{cum}}\\ & u_t \geq u_{t-1}\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

Mean-Max-DD portfolio as an LP

- We can finally write the maximization of the mean return subject to the Max-DD constraint as

$$
\begin{array}{ll}
\underset{\mathbf{w},\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\
\text{subject to} & \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq u_t \leq \mathbf{w}^T\mathbf{r}_t^{\text{cum}} + c, \quad \forall 1 \leq t \leq T \\
& u_{t-1} \leq u_t \\
& \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.
\end{array}
$$

Mean-Ave-DD portfolio as an LP

- Similarly, we can consider the maximization of the mean return subject to an Ave-DD constraint:

$$\begin{array}{ll}\underset{\mathbf{w},\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \frac{1}{T}\sum_{t=1}^{T}(u_t - \mathbf{w}^T\mathbf{r}_t^{\text{cum}}) \leq c\\ & u_t \geq \mathbf{w}^T\mathbf{r}_t^{\text{cum}}\\ & u_{t-1} \leq u_t\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}\end{array}$$

or, equivalently,

$$\begin{array}{ll}\underset{\mathbf{w},\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \frac{1}{T}\sum_{t=1}^{T} u_t \leq \sum_{t=1}^{T}\mathbf{w}^T\mathbf{r}_t^{\text{cum}} + c\\ & \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq u_t\\ & u_{t-1} \leq u_t\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

Mean-CDaR portfolio as an LP

- Finally, we can consider the maximization of the mean return subject to a CDaR constraint:

$$\begin{array}{ll}\underset{\mathbf{w},\zeta}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T}\left[\max_{1\leq\tau\leq t}\mathbf{w}^T\mathbf{r}_{\tau}^{\text{cum}} - \mathbf{w}^T\mathbf{r}_{t}^{\text{cum}} - \zeta\right]^{+} \leq c\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

- Similarly to the CVaR case, we can get rid of the $[\cdot]^{+}$ operator by introducing some additional variables:

$$\begin{array}{ll}\underset{\mathbf{w},\{z_t\},\zeta}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu}\\ \text{subject to} & \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T} z_t \leq c\\ & 0 \leq z_t \geq \max_{1\leq\tau\leq t}\mathbf{w}^T\mathbf{r}_{\tau}^{\text{cum}} - \mathbf{w}^T\mathbf{r}_{t}^{\text{cum}} - \zeta, \quad t = 1,\dots,T\\ & \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.\end{array}$$

Mean-CDaR portfolio as an LP

- Similarly to the Max-DD and Ave-DD cases, we can get rid of the max operators by introducing additional variables:

$$
\begin{array}{ll}
\underset{\mathbf{w},\{z_t\},\zeta,\{u_t\}}{\text{maximize}} & \mathbf{w}^T\boldsymbol{\mu} \\
\text{subject to} & \zeta + \frac{1}{1-\alpha}\frac{1}{T}\sum_{t=1}^{T} z_t \leq c \\
& 0 \leq z_t \geq u_t - \mathbf{w}^T\mathbf{r}_t^{\text{cum}} - \zeta, \quad t = 1,\dots,T \\
& \mathbf{w}^T\mathbf{r}_t^{\text{cum}} \leq u_t \\
& u_{t-1} \leq u_t \\
& \mathbf{1}^T\mathbf{w} = 1, \quad \mathbf{w} \geq \mathbf{0}.
\end{array}
$$

Word of caution on DD

- The maximum drawdown is extremely sensitive to minute changes in the portfolio weights and to the specific time period examined.
- If returns are close to normally distributed, the distribution of drawdowns is just a function of the variance, so there's no need to include drawdowns explicitly in your portfolio construction objective. Minimizing variance is the same as minimizing expected drawdowns.
- On the other hand, if returns are very non-normal and you want to find a portfolio that minimizes the expected drawdowns, you still wouldn't choose weights that minimize historical drawdown. Why?
- Because minimizing historical drawdown is effectively the same as taking all your returns that weren't part of a drawdown, and hiding them from your optimizer, which will lead to portfolio weights that are a lot less accurately estimated than if you let your optimizer see all the data you have.
- Instead, you might just include terms in your optimization objective that penalize negative skew and penalize positive kurtosis.

Outline

Conclusions

- We have reviewed the Markowitz portfolio formulation and understood that it has many practical flaws that make it impractical. Indeed, it is not used by practitioners.
- We have learned about alternative measures of risk as opposed to variance:
  - Downside risk (one particular example is the semi variance)
  - VaR
  - CVaR
  - Drawdown
- We have formulated several portfolio designs based on downside risk, CVaR, and DD, all as LPs!

Thanks

For more information visit:

https://www.danielppalomar.com