Data-driven robust mean-CVaR portfolio selection under distribution ambiguity

Firoozeh Abrishami & Hossein Rahmani

Sharif University of Technology

Financial Mathematics
July 18, 2023

Presentation Overview

1. Introduction
2. Notation
3. Mean CVaR Optimization
4. Calibration of γ_1 and γ_2

Introduction

- Move from class problem to minimization of worst-case CVaR
- Optimization over ambiguity set reformulated as SOCP
- why CVaR over VaR? (coherency)
- Historical data, Expert Knowledge, or Ambiguity set?

Notation

S_+^n denotes the cone of positive semidefinite matrices. For $X \in S_+^n$, we denote its symmetric square root by $\sqrt(X)$ or $X^{\frac{1}{2}}$ .For a given reversible matrix $G \in S^n$, $||z||_G$ represents the ellipsoidal norm of a vector z, i.e. $||z||_G = \sqrt{z^T G^{-1} z}$.
For a matrix $X \in S^n$, $||X||_F$ represents the Frobenius norm, i.e. $||X||_F = (X \circ X)^{\frac{1}{2}} = \sqrt{tr(XX^T)}$.

Basic Definition

$$VaR_\beta(x) = \min\{\eta \in R : \int_{\xi : l(x,\xi) \leq \eta} p(\xi) d\xi \geq \beta\}. \quad (1)$$

Our intuition about it is that the minimum amount of loss is when we have a total return of at least β.

CVaR

CVaR at level $1-\beta$ with respect to the distribution P, which is defined as the expected value of the loss $l(x,\xi)$ exceeding VaR, can be expressed as:

$$CVaR_\beta(x,P) = E_P[l(x,\xi)|l(x,\xi) \geq VaR_\beta(x)] \tag{2}$$

$$= \frac{1}{1-\beta}\int_{\xi:l(x,\xi)\geq VaR(x)} l(x,\xi)p(\xi)d\xi. \tag{3}$$

Another Formulation

By theory from [Rockafellar and Uryasev (2000)], we know that the calculation of CVaR can be achieved by minimizing the following function

$$F_\beta(x, \eta) = \eta + \frac{1}{1-\beta} \int_{\xi \in R^n} [l(x, \xi) - \eta]^+ p(\xi) d\xi \qquad (4)$$

where $[t]^+ = \max\{0, t\}$. So, we have this equation $CVaR_\beta(x, P) = \min_{\eta \in R}\ F_\beta(x, \eta)$.

WCVaR

Definition 1:
Given a probability threshold $\beta > 0$, the worst-case CVaR (WCVaR) of portfolio x, where random vector ξ may assume a distribution from ambiguity set D, is defined by

$$WCVaR_\beta(x) = \sup_{P \in D} CVaR_\beta(x, P). \tag{5}$$

WCVaR inherits subadditivity, positive homogeneity, monotonicity and translation invariance. Therefore, WCVaR, like CVaR, is a coherent risk measure.

Our Problem

Consider a financial market consisting of n different assets. A portfolio is characterized by a vector of asset weights $x \in R^n$ (adds up to 1).It means that the component x_i denotes the percentage of total wealth that is invested in the *i*th asset at the beginning of the investment period

Classic CVaR (MC) problem is this

$$\min_x CVaR_\beta(x, P), \tag{6}$$

$$s.t. \;\; E_P(\xi)^T x \geq \rho, \;\; x \in X \tag{7}$$

where ρ stands for the lower limit on the target expected return.

SMC

One approach is sample-based mean-CVaR optimization problem (SMC)

$$(SMC): \min_{(x,\eta)} \hat{F}_\beta(x,\eta), \tag{8}$$

$$s.t. \ \frac{1}{S}\sum_{k=1}^{S}(\xi[k])^T x \geq \rho \tag{9}$$

$$x \in X, \eta \in R \tag{10}$$

SMC is easy to solve if X is convex and l(x,) is convex in x (Rockafellar and Uryasev 2000). and depends on small portion of points but simultaneously its result is unreliable.

Ambiguity Set

$$D_F\{\gamma_1, \gamma_2\} = \Bigg\{ P \in M_+ :$$

$$\mathrm{P}(\xi \in \Omega) = 1, (E_P(\xi) - \hat{\mu})^T \hat{\Sigma}^{-1} (E_P(\xi) - \hat{\mu}) \leq \gamma_1, (11)$$

$$||Cov_P(\xi)\hat{\Sigma}||_F \leq \gamma_2, \quad Cov_P(\xi) > 0 \qquad (12)$$

$$\hat{\Sigma} = \frac{1}{S-1} \sum_{i=1}^{S} (\xi^{[i]} - \hat{\mu})(\xi^{[i]})^T \qquad (13)$$

$$\hat{\mu} = \frac{1}{S} \sum_{i=1}^{S} \xi^{[i]} \qquad (14)$$

lemma

Lemma 2.1: [Chen et al. 2011] Assume that $l(x,\xi) = -\xi^T x$ and random vector $\xi \in R^n$, with mean $\hat{\mu}$ and covariance $\hat{\Sigma} > 0$, follows a family of distributions F, which is defined by $F = \{P \in M^+ | P(\xi \in \Omega) = 1, E_P(\xi) = \hat{\mu}, Cov_P(\xi) = \hat{\Sigma}\}$. If the support set of ξ covers the whole space, i.e. $\Omega = R^n$, then we have

$$\max_{P \in F} CVaR_\beta(x, P) = -\hat{\mu}^T x + \kappa \sqrt{x^T \hat{\Sigma} x}, \tag{15}$$

where $\kappa = \sqrt{\frac{\beta}{1-\beta}}$.

The details of the proof of Lemma 2.1 are referred to [Chen et al. 2011].

Our Optimization

$$
\begin{aligned}
&\min_{x,s,t} \kappa s - \hat{\mu}^T x + \sqrt{\gamma_1} t, && (16)\\
s.t.\ &\sqrt{\gamma_1} ||\hat{\Sigma}^{\frac{1}{2}} x||_2 \leq \hat{\mu}^T x - \rho, && (17)\\
&||(\hat{\Sigma} + \gamma_2 I_n)^{\frac{1}{2}} x||_2 \leq s, && (18)\\
&||\hat{\Sigma}^{\frac{1}{2}} x||_2 \leq t, && (19)\\
&x \in X && (20)
\end{aligned}
$$

variables are $x \in R^n,\ s, t \in R$ and it's a SOCP.

Zero-Net adjustment

we have an extra condition $e^T(E_P[\xi] - \hat{\mu}) = 0$ in the ambiguity set.

$$D_F^{adj}\{\gamma_1, \gamma_2\} = \left\{ P \in M_+ : \right. \quad (21)$$

$$P(\xi \in \Omega) = 1, (E_P(\xi) - \hat{\mu})^T \hat{\Sigma}^{-1} (E_P(\xi) - \hat{\mu}) \leq \gamma_1, \quad (22)$$

$$||Cov_P(\xi)\hat{\Sigma}||_F \leq \gamma_2, \ \ Cov_P(\xi) > 0 \quad (23)$$

$$\left. e^T(E_P(\xi) - \hat{\mu}) = 0 \right\} \quad (24)$$

Zero-Net adjustment

Our problem changes to

$$
\begin{aligned}
\min_{x,s,t} \; & \kappa\sigma - \hat{\mu}^T x + \sqrt{\gamma_1}\omega, && (25)\\
s.t. \;\; & \sqrt{\gamma_1}||\Lambda^{\frac{1}{2}} x||_2 \leq \hat{\mu}^T x - \rho, && (26)\\
& ||(\hat{\Sigma} + \gamma_2 I_n)^{\frac{1}{2}} x||_2 \leq \sigma, && (27)\\
& ||\Lambda^{\frac{1}{2}} x||_2 \leq \omega, && (28)\\
& x \in X && (29)
\end{aligned}
$$

Our variables are $x \in R^n, \;, s, t \in R \;\; \Lambda = \hat{\Sigma} - \frac{1}{e^T\hat{\Sigma}e}\hat{\Sigma}ee^T\hat{\Sigma}$

bootstrap

Bootstrapping procedure:

Step 1. Construct B bootstrap samples $\{Y_1, Y_2, \ldots, Y_B\}$ (for example, $B = 10000$) by drawing random observations with replacement from the available observations.

Step 2. For each bootstrap sample Y_b, compute the corresponding mean $\hat{\mu}_b$ and covariance matrix $\hat{\Sigma}_b$, and then generate a sample

$$C = \left\{(\hat{\mu}_b, \hat{\Sigma}_b) : b = 1, \ldots, B\right\}.$$

Step 3. For sample C, define data sets C_{γ_1} and C_{γ_2} as

$$C_{\gamma_1} = \left\{\gamma_{1b} : \gamma_{1b} = (\hat{\mu}_b - \hat{\mu})^{\mathrm{T}} \hat{\Sigma}^{-1} (\hat{\mu}_b - \hat{\mu}), \right.$$

$$\left. b = 1, \ldots, B\right\},$$

$$C_{\gamma_2} = \left\{\gamma_{2b} : \gamma_{2b} = \|\hat{\Sigma}_b - \hat{\Sigma}\|_F, b = 1, \ldots, B\right\}$$

to ensure reasonable values of γ_1 and γ_2. The percentiles of the empirical distributions of C_{γ_1} and C_{γ_1} can then be referenced to derive γ_1 and γ_2. Consequently, the calibrated values of γ_1 and γ_2 are

$$\hat{\gamma}_1 = q_\zeta(C_{\gamma_1}), \quad \hat{\gamma}_2 = q_\zeta(C_{\gamma_2}),$$

where $q_\zeta(\cdot)$ is an upper quantile of the corresponding data sets (for example, $\zeta = 95\%$).