Robust Stochastic Portfolio Optimization: A Data-driven Clustering Approach

清华大学
大学生学术研究推进论坛
Tsinghua Student Research Conference

Investigator: Tianyu Wang, SEM. Advisor: Chun Wang, SEM. Email: *wangty.17@sem.tsinghua.edu.cn*

Keywords: Robust Optimization, Portfolio Control, CVaR

Introduction

Optimal portfolio selection has received attention since Markowitz (1952). And one important criterion among this is to **minimize CVaR (conditional value-at-risk)**. However in real problems future returns vary greatly due to uncertainty. **Distributionally robust optimization** is to optimize out-of-sample performance by estimated worst-case distribution. Our work intends to gain **mean-covariance information** by clustering from history data, thus can remain computational tractability in real datasets with **side information**.

Problem Formulations

Notations x decision weight.

r random variable for assets return.

s different scenarios.

(μ_s, Σ_s) mean-covariance for each scenario

Model Clustering information can easily be done by K-Means. Consider the **ambiguity sets**

$$\mathcal{F}(\mu, \Sigma) = \left\{ \mathbb{P} \in \mathcal{P}_0\left(\mathbb{R}^{I_r} \times [S]\right) \middle| \begin{array}{ll} (\tilde{r}, \tilde{s}) \sim \mathbb{P} & \\ \mathbb{E}_{\mathbb{P}}[\tilde{r}_s] = \mu_s & \forall s \in [S] \\ \mathbb{E}_{\mathbb{P}}\left[(\tilde{r} - \mu_s)(\tilde{r} - \mu_s)^{\top}\right] = \Sigma_s & \forall s \in [S] \\ \mathbb{P}[\tilde{s} = s] = p_s & \forall s \in [S] \end{array} \right\}$$

As the curse of dimensionality of K-means, we further choose **Fama-French three-factor as side information** for clustering and compute values.

Target By minimizing worst-case CVaR model under distribution F we can realize optimal portfolio selection:

$$\inf_{x,v} \left\{ v + \frac{1}{\epsilon} \sup_{\mathbb{P} \in \mathbb{F}} \mathbb{E}_{\mathbb{P}}\left[(-\tilde{r}'x - v)^+\right] \right\}$$

Similar transformations like Popescu (2007) can convert it to the following:

$$\inf_{x,v} \left\{ v + \frac{1}{2\epsilon} \sum_{s=1}^{K} p_s(-\mu_s' x - v + \sqrt{x'\Sigma_s x + (\mu_s' x + v)^2}) \right\}$$

It reduces conservativeness and can change to **SOCP,** guaranteeing its computational tractability.

Benchmark Below are other policies for evaluation:

Method	Target	Parameter
1/N Policy	$x = (\frac{1}{N}, \frac{1}{N}, ..., \frac{1}{N})^T$	NA
Markowitz	$\sup_x \{\mu'x - \frac{\gamma}{2} x'\Sigma x\}$	$\gamma = 0.5$
CVaR (SAA)	$\inf_{x,v} \left\{ v + \frac{1}{\epsilon}\frac{1}{M} \sum_{i \in [M]} (-r_i' x - v)^+ \right\}$	$\epsilon = 0.05$
F-CVaR (Popescu)	$\inf_{x,v} \{ v + \frac{1}{2\epsilon}[(-\mu'x - v) + \sqrt{x'\Sigma x + (\mu'x + v)^2}] \}$	$\epsilon = 0.05$

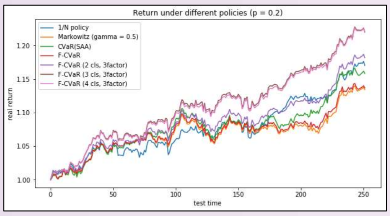
Computational Experiments

Method	Sharpe ratio	VaR	CVaR
1/N Policy	0.1460	0.5560	0.0481
Markowitz (0.5)	0.1226	0.5614	0.0397
CVaR (SAA)	0.1454	0.6198	0.0397
F-CVaR (Popescu)	0.1304	0.6112	0.0392
F-CVaR (2 cls, 3 factor)	0.1773	0.5306	0.0416
F-CVaR (3 cls, 3 factor)	0.1935	0.4891	0.0461
F-CVaR (4 cls, 3 factor)	0.1953	0.4799	0.0460
F-CVaR (2 cls, return)	0.1692	0.5263	0.0409
F-CVaR (3 cls, return)	0.1589	0.5571	0.0453
F-CVaR (4 cls, return)	0.1531	0.5639	0.0473

Datasets
10 Industry
2012 - 2017
Method
rolling window
5 days
Criteria
Sharpe ratio
VaR / CVaR

Clustering approach especially three-factor performs better compared with benchmarks**.**

Extensions and Conclusion



(p = 0.2 is the barrier factor for rebalancing costs, multiplying $1 - p\sum_{j=1}^{N} |x_{j,t+1} - x_{j,t^+}|$ in return)

Extensions
Even if we consider transaction costs in, clustering approach performs better.

Conclusion Compared with other DRO approaches, clustering approach is more **tractable**. Moreover, it can form the framework for various **side information.**

Traditional DRO

Assets return

Side information (factor/market...)

Clustering Approach

Acknowledgement and References

The mathematical programs in numerical experiments are solved using Gurobl 8.1.0 in Python 3.7. The datasets are from the Kenneth R. French – Data Library website and the work is supervised by Melvyn Sim from NUS too.

Markowitz, H. 1952. Portfolio selection. J. Finance 7(1) 77–91.

Popescu, Ioana. 2007. Robust mean-covariance solutions for stochastic optimization. Operations Research 55(1) 98 - 112.