FE670 Algorithmic Trading Strategies

Lecture 8. Robust Portfolio Optimization

Steve Yang

Stevens Institute of Technology

steve.yang@stevens.edu

10/26/2023

Outline

Robust Mean-Variance Formulations

- Uncertainty in the inputs to a portfolio optimization problem (for example, the expected returns and their variances and covariances) can be modeled directly in the optimization process.

- We recall that the classical mean-variance problem introduced earlier

$$\max_{w} \boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

$$s.t. \mathbf{w}'\mathbf{l} = 1$$

where $\mathbf{l} = [1, 1, .., 1]'$. In this optimization problem $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda$, and $\mathbf{w}$ denote the expected return, asset return covariance matrix, risk aversion coefficient, and portfolio weights, respectively.

Uncertain in Expected Return Estimates

- An easy way to incorporate uncertainty caused by estimation errors is to require that the investor be protected if the estimated expected return $\hat{\boldsymbol{\mu}}_i$ for each asset is around the true expected return μ_i. The error from the estimation can be assumed to be not larger than some small number $\delta_i > 0$. A simple choice for the *uncertainty set* for $\boldsymbol{\mu}$ is

$$U_\delta(\hat{\boldsymbol{\mu}}_i) = \{\boldsymbol{\mu} | |\mu_i - \hat{\mu}_i| \leq \delta_i, i = 1, ..., N\}$$

  where $\mathbf{l} = [1, 1, .., 1]'$. In this optimization problem $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda$, and $\mathbf{w}$ denote the expected return, asset return covariance matrix, risk aversion coefficient, and portfolio weights, respectively.

  The δ_i's could be specified by assuming some confidence interval around the estimated expected return.

- The robust formulation of the mean-variance problem under the preceding assumption on $\hat{\mu}_i$ is

$$\max_{w} \boldsymbol{\mu}'\mathbf{w} - \boldsymbol{\delta}'|\mathbf{w}| - \lambda\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$
$$s.t.\mathbf{w}'\mathbf{l} = 1$$

- If the weight of asset i in the portfolio is negative, the worst-case expected return for asset i is $\mu_i + \delta_i$ (we lose the largest amount possible). If the weight of asset i in the portfolio is positive, then the worst-case expected return for asset i is $\mu_i - \delta_i$ (we gain smallest amount possible).
- The objective agrees with our intuition: it tries to minimize the worst-case expected portfolio return. Assets whose mean return estimates are less accurate (have a larger estimation error δ_i) are penalized in the objective function, and will tend to have smaller weights in the optimal portfolio allocation.

To gain some additional insight, let us rewrite the robust formulation as

$$\max_{w}(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\delta,\mathbf{w}})\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

$$s.t. \mathbf{w}'\mathbf{I} = 1$$

where

$$\boldsymbol{\mu}_{\delta,\mathbf{w}}(\mathbf{w}) = \begin{bmatrix} \text{sign}(w_1)\delta_1 \\ \text{sign}(w_2)\delta_2 \\ \vdots \\ \text{sign}(w_N)\delta_N \end{bmatrix}$$

Here $\text{sign}(\cdot)$ is the sign function (that is, $sign(x) = 1$ when $x \geq 0$ and $sign(x) = -1$ when $x < 0$). In this reformulation of the problem we see that robust optimization is related to statistical shrinkage, and the original expected return vector is shrunk to $\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\delta,\mathbf{w}}$.

By using the equality $w_i \text{sign}(w_i)\delta_i = w_i \frac{w_i}{|w_i|}\delta_i = \frac{w_i}{\sqrt{|w_i|}}\delta_i \frac{w_i}{\sqrt{|w_i|}}$, we can rewrite the problem as

$$\max_{w} \boldsymbol{\mu}'\mathbf{w} - \lambda\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \hat{\mathbf{w}}'\boldsymbol{\Delta}\hat{\mathbf{w}}$$
$$s.t.\mathbf{w}'\mathbf{I} = 1$$

where

$$\hat{\mathbf{w}} = \begin{bmatrix} \frac{w_1}{\sqrt{|w_1|}} \\ \vdots \\ \frac{w_N}{\sqrt{|w_N|}} \end{bmatrix} \boldsymbol{\Delta} = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_N \end{bmatrix}$$

Observe that this problem is yet another modification of the classical mean-variance problem. In particular, a *risk-like* term $\hat{\mathbf{w}}'\boldsymbol{\Delta}\hat{\mathbf{w}}$ has been added to the classical formulation. This term can be interpreted as a risk adjustment performed by an investor who is averse to estimation error.

- One can define many other uncertainty sets for the expected returns vector $\boldsymbol{\mu}$. While more general uncertainty sets lead to more complicated optimization problems, the basic intuition and interpretation remain the same. For instance, consider the uncertainty set:

$$U_\delta(\hat{\boldsymbol{\mu}}_i) = \left\{(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}})'\boldsymbol{\Sigma}_\mu^{-1}(\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}) \leq \delta^2\right\}$$

  It captures the idea that the investor would like to be protected in instances in which the total scaled deviation of the realized average returns from the estimated returns is within δ.
- We may ask ourselves what the worst estimates of the expected returns would be, and how we would allocate the portfolio in this case. Mathematically, this can be expressed as:

$$\max_{w} \min_{\boldsymbol{\mu}\in\left\{(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})'\boldsymbol{\Sigma}_\mu^{-1}(\boldsymbol{\mu}-\hat{\boldsymbol{\mu}})\leq\delta^2\right\}} [\boldsymbol{\mu}'\mathbf{w} - \lambda\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}]$$

$$s.t. \mathbf{w}'\mathbf{I} = 1$$

- This problem is called the *robust counterpart*, or the *max-min* problem, and it is not in a form that can be put into a standard optimization solver.
- We need to solve the inner problem first while holding the vector of weights $\mathbf{w}$ fixed, and compute the worst expected portfolio return over the set of possible values for μ:

$$\min_{\mu} \boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$
$$\text{s.t. } (\mu - \hat{\mu})'\boldsymbol{\Sigma}_{\boldsymbol{\mu}}^{-\mathbf{1}}(\mu - \hat{\mu}) \leq \delta^2$$

  The lagrangian of this problem takes the form $L(\boldsymbol{\mu}, \gamma) = \boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \gamma(\delta^2 - (\mu - \hat{\mu})'\boldsymbol{\Sigma}_{\boldsymbol{\mu}}^{-\mathbf{1}}(\mu - \hat{\mu}))$.
- Differentiating this with respect to $\boldsymbol{\mu}$, we obtain the first-order condition and the optimal value is

$$\boldsymbol{\mu}^* = \hat{\boldsymbol{\mu}} - \frac{1}{2\gamma}\boldsymbol{\Sigma}_{\mu}\mathbf{w}$$

- The optimal value of γ can be found by maximizing the Lagrangian after substituting the expression for the worst-case μ, that is

$$\max_{\gamma \geq 0} L(\boldsymbol{\mu}, \gamma) = \boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \frac{1}{4\gamma}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \gamma\delta^2$$

After solving the first-order condition, we obtain

$$\boldsymbol{\gamma}^* = \frac{1}{2\delta}\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\mu \mathbf{w}}$$

Finally, by substituting the expression for $\gamma*$ in the Lagrangian, we obtain the robust problem

$$\max_{w} \boldsymbol{\mu}'\mathbf{w} - \lambda \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w} - \delta\sqrt{\mathbf{w}'\boldsymbol{\Sigma}_\mu \mathbf{w}}$$

$$s.t. \mathbf{w}'\mathbf{I} = 1$$

- In fact, critics of the approach have argued that the realized returns typically have large stochastic components that dwarf the expected returns, and hence estimating $\boldsymbol{\Sigma}_\mu$ accurately from historical data is very hard, if not possible [Lee, Stefek and Zhelenyak, 2006].

- Several approximation methods for estimating $\boldsymbol{\Sigma}_\mu$ have been found to work well in practice. For example, using just the diagonal matrix containing the variances of the estimates frequently provide most of the benefit in robust portfolio optimization [Stubbs and Vance, 2005].

- Other effective estimation error covariance matrix include least squares regression models, the James-Stein estimator, and the Black-Litterman model.

Least Squares Regression Models

- If expected returns are estimated based on linear regression, then one can calculate an estimate of the error covariance matrix from the regression errors. Let us assume we have the factor model for the returns:

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

  This can be written as

  $$\mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \boldsymbol{\epsilon}_i \text{ or } \mathbf{Y} = \mathbf{A}\mathbf{X} + \boldsymbol{\epsilon}$$

- If a portfolio manager decomposes the expected return forecast into factor-specific and asset specific returns, then he is concerned about the standard error covariance matrix for the intercept term $\boldsymbol{\mu}$. The matrix of estimation errors $\boldsymbol{\Sigma}_{\mu}$ for the response corresponding to the factor realizations $\mathbf{f}_{\tau} = (1, \mathbf{f}_{\tau 1}, ..., \mathbf{f}_{\tau M})' \in \mathbf{R}^M$ is given by:

  $$\mathbf{f}'_{\tau}\mathbf{X}\mathbf{X}'^{-1}\mathbf{f}_{\tau}\left\{\frac{1}{T}(\mathbf{Y} - \mathbf{A}\mathbf{X})'(\mathbf{Y} - \mathbf{A}\mathbf{X})\right\}$$

The James-Stein Estimator

- The James-Stein estimator of expected returns is computed as a weighted average of the sample average returns (computed from a sample of size T) and a shrinkage target of $\boldsymbol{\mu}_0$.

$$\hat{\boldsymbol{\mu}}_{JS} = (1 - w)\hat{\boldsymbol{\mu}} + w\boldsymbol{\mu}_0$$

  The special form of the James-Stein shrinkage estimator proposed by Jorion (named the Bayes-Stein estimator) is based on Bayesian methodology. The shrinkage target $\boldsymbol{\mu}_0$ is computed as

$$\boldsymbol{\mu}_0 = \frac{\mathbf{l}'\boldsymbol{\Sigma}^{-1}}{\mathbf{l}'\boldsymbol{\Sigma}^{-1}\mathbf{l}}\hat{\boldsymbol{\mu}}$$

  where $\boldsymbol{\Sigma}$ is the real covariance matrix of the N returns. This matrix is unknown in practice, but one can replace Σ in the previous equation by (where, S is the usual sample covariance matrix.):

$$\hat{\boldsymbol{\Sigma}} = \frac{T-1}{T-N-3}\mathbf{S}$$

The James-Stein Estimator

- The variance of the Bayes-Stein estimator for the expected returns is given by

$$var(\hat{\boldsymbol{\mu}}_{BS}) = \boldsymbol{\Sigma} + \frac{1}{T+\tau}\boldsymbol{\Sigma} + \frac{\tau}{T(T+\tau+1)}\frac{\boldsymbol{\mu}'}{\boldsymbol{l}'\boldsymbol{\Sigma}\boldsymbol{l}}$$

  and can be used as an estimate for the error covariance matrix Σ_μ. The parameter τ is a scalar that describes the confidence in the precision of estimation of the covariance matrix $\boldsymbol{\Sigma}$.

The Black-Litterman Model

- The Black-Litterman model for estimating expected returns combines the market equilibrium with an investor's views. The formula for the estimate is

$$\begin{aligned}\hat{\mu}_{BL} &= (\mathbf{X}^{'}\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \\ &= [(\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}[(\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1}\boldsymbol{\Pi} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{q}]\end{aligned}$$

  where $\boldsymbol{\Sigma}$ is the covariance matrix of returns; $\boldsymbol{\Pi}$ is the vector of expected excess returns, computed from an equilibrium model such as CAPM; $\boldsymbol{\tau}$ is a scalar that represents the confidence in the estimation of the market prior; $\mathbf{q}$ is vector of investor's views; $\mathbf{P}$ is a matrix of investor's views; $\boldsymbol{\Omega}$ is matrix expressing the confidence in the investor's views.

- The covariance of the expected return is $[(\boldsymbol{\tau}\boldsymbol{\Sigma})^{-1} + \mathbf{P}'\boldsymbol{\Omega}^{-1}\mathbf{P}]^{-1}]$. It can be used an approximation from the estimation error covariance matrix $\boldsymbol{\Sigma}_{\mu}$.

Uncertainty in Return Covariance Matrix Estimates

Mean-Variance portfolio optimization is less sensitive to inaccuracies in the estimate of the covariance matrix $\boldsymbol{\Sigma}$ than it is to estimation errors in expected returns.

$$\max_{w} \left\{ \min_{\mu \in U_\mu} \{\boldsymbol{\mu}'\mathbf{w}\} - \lambda \max_{\Sigma \in U_\Sigma} \{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\} \right\}$$
$$s.t. \mathbf{w}'\mathbf{I} = 1$$

where U_μ and U_Σ denote the uncertainty sets of expected returns and covariances, respectively.

- A few different methods for modeling uncertainty in the covariance matrix are used in practice. Some are superimposed on top of factor models for returns, while others consider confidence intervals for the individual covariance matrix entries.

Factor Models

- If we assume a standard factor model for returns

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

Then the covariance matrix of returns $\boldsymbol{\Sigma}$ can be expressed as

$$\boldsymbol{\Sigma} = \mathbf{V}'\mathbf{F}\mathbf{V} + \mathbf{D}$$

It is assumed that the vector of residual returns $\boldsymbol{\epsilon}$ is independent of the vector of factor returns $\mathbf{f}$ and that the variance of $\boldsymbol{\mu}$ is zero.

- The statistical properties of the estimate of $\mathbf{V}$ naturally lead to an uncertainty set of the kind.

$$S_v = \{\mathbf{V} : \mathbf{V} = \mathbf{V}_0 + \mathbf{W}, ||\mathbf{W}_i|| \leq \rho_i, i = 1, ..., N\}$$

where $\mathbf{W}_i$ denotes the i-th column of $\mathbf{W}$ and $||\mathbf{W}_i|| = \sqrt{\mathbf{w}'\mathbf{G}\mathbf{w}}$. (The definition of the matrix $\mathbf{G}$ can be based on the data used to produce the estimates of the regression coefficients of the factor model.)

Confidence Intervals for the Entries of the Covariance Matrix

- Instead of using uncertainty sets based on estimates from a factor model, one can specify intervals from the individual elements of the covariance matrix of the kind

$$\underline{\boldsymbol{\Sigma}} \leq \boldsymbol{\Sigma} \leq \overline{\boldsymbol{\Sigma}}$$

If we assume that the estimates of expected returns vary in intervals

$$U_{\delta}(\hat{\boldsymbol{\mu}}_i) = \{\boldsymbol{\mu} | |\mu_i - \hat{\mu}_i| \leq \delta_i, i = 1, ..., N\}$$

short sales are not allowed (i.e., $\mathbf{w} \geq 0$, and the matrix $\overline{\boldsymbol{\Sigma}}$ is positive semidefinite, the resulting optimization problem is very simple to formulate.

$$\max_{w} \left\{ \min_{\mu \in U_{\mu}} \{\boldsymbol{\mu}'\mathbf{w}\} - \lambda \max_{\Sigma \in U_{\Sigma}} \{\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}\} \right\}$$

$$s.t. \mathbf{w}'\mathbf{I} = 1$$

Empirical Comparison of Mean-Variance vs. Robust Optimization

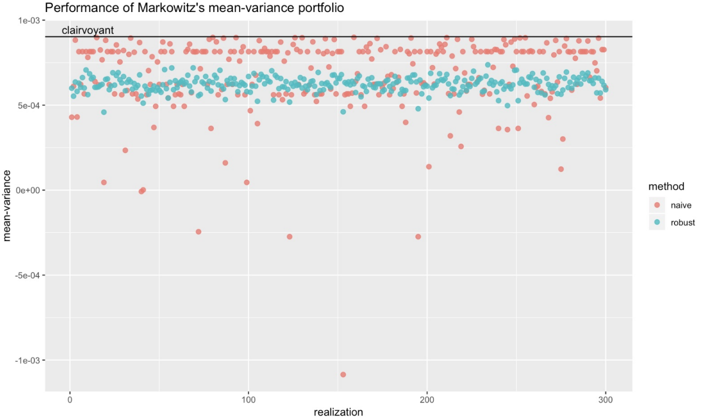

Figure: The naive may achieve a better mean-variance objective than the robust but occasionally can be pretty bad. The robust is more stable. HKUST

Empirical Comparison of Mean-Variance vs. Robust Optimization

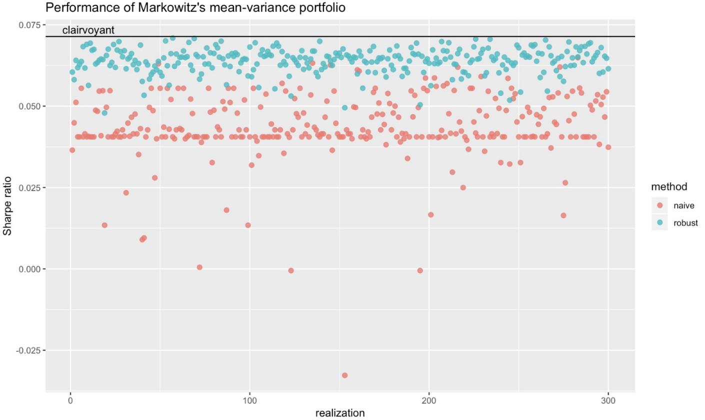

Figure: In terms of Sharpe ratio, the robust is clearly superior to the naive (note that the mean-variance portfolio is not the same as the maximum Sharpe ration portfolio). HKUST

Using Robust Mean-Variance Portfolio Optimization Example in Python

- The robust model is presented below (Bertsimas and Sim, 2004).

$$\max_{w}\left\{\min_{\mu\in Z_{\mu}}\{\mathbf{\Sigma}_{i=1}^{n}(p_i+\delta_i z_i)w_i)\}\right\}$$

$$s.t.\mathbf{w}'\mathbf{I}=1$$

- where the affine term $p_i+\delta_i z_i$ represents the random stock return, and the random variable is between $[-1,1]$, and the stock return has an arbitrary distribution in the interval $[p_i-\delta_i, p_i+\delta_i]$
- The uncertainty set $\boldsymbol{Z}_\mu$ is given as $Z_\mu=\{z_i : ||z_i||_\infty \leq 1, ||z_i||_1 \leq \Gamma\}$, where Γ is the budget of uncertainty parameter.

Using Robust Mean-Variance Portfolio Optimization Example in Python

- Values of the budget of uncertainty and other parameters are presented as follows:

$$\Gamma = 5$$
$$p_i = 1.15 + i\frac{0.05}{150}, i = 1, ..., n$$
$$\delta_i = \frac{0.05}{450}\sqrt{2 * i * n(n+1)}, i = 1, ..., n$$

- The robust optimization model can be implemented by the Python code.

Using Robust Mean-Variance Portfolio Optimization in Practice

- Robust optimization does, however, come at the cost of additional modeling effort. The important question is whether this effort is worthwhile.
- Critics have argued that robust optimization does not provide more benefit than shrinkage estimators that combine the minimum variance portfolio with a speculative investment portfolio. However, robust optimization is not necessarily equivalent to shrinkage estimation. They are particularly different in the presence of additional portfolio constraints.
- An advantage of the robust optimization approach is that the parameter values in the robust formulation can be matched to probabilistic guarantees.

- The discussion in the previous sections leads to the question: so which approach is best for modeling financial portfolios?

  The short answer is: it depends. It depends on the size of the portfolio, the type of assets and their distributional characteristics, the portfolio strategies and trading styles involved, and exiting technical and intellectual infrastructure, among others.

  Sometimes it makes sense to combine several techniques, such as a blend of Bayesian estimation and robust portfolio optimization. This is an empirical question; indeed, the only way to find out is through extensive research and testing.

- A simple check-list for robust quantitative portfolio management:
    1. Risk forecasting: develop an accurate risk model.
    2. Return forecasting: construct robust expected return estimates.
    3. Classical protfolio optimization: start with a simple framework.
    4. Mitigate model risk:
        a. Minimize estimation risk through the use of robust estimators.
        b. Improve the stability of the optimization framework through robust optimization.
    5. Extensions.

  In general, the most difficult item in this list is to calculate robust expected return estimates. Developing profitable trading strategies is hard, but not impossible. It is important to remember that modern portfolio optimization techniques and fancy mathematics are not going to help much if the underlying trading strategies are subpar.