Robust Portfolio Optimization

November 2007

Prof. Dr. Heinz Zimmermann
Wirtschaftswissenschaftliches Zentrum (WWZ), Universität Basel
heinz.zimmermann@unibas.ch

Daniel Nierdermayer
Wirtschaftswissenschaftliches Zentrum (WWZ), Universität Basel
daniel.niedermayer@vwi.unibe.ch
Overview

1. Estimation Error
2. Robust Optimization
3. Reverse Optimization
1. Estimation Error

2. Robust Optimization

3. Reverse Optimization
1. Estimation Error

a. Monte Carlo
b. Experiment
c. Results

2. Robust Optimization

3. Reverse Optimization
Traditional Mean-Variance optimization (Markowitz) assumes known expected returns $\bar{\mu}$ and covariance matrix $\Sigma$.

In practice, $\bar{\mu}$ and $\Sigma$ must be estimated and therefore contain estimation error.

This section presents a simulation that highlights the impact of estimation error on optimal asset allocation.

As the simulation uses multivariate normal returns, we first discuss a return generating Monte Carlo method.
Overview

1. Estimation Error
   a. Monte Carlo
   b. Experiment
   c. Results

2. Robust Optimization

3. Reverse Optimization

a. Generating multivariate normal returns (Monte Carlo Simulation)

b. Experiment
Let $x$ be an $N \times 1$ vector with

$$x \sim \mathcal{N}(0, I_{N \times N}),$$

where $I_{N \times N}$ is an $N \times N$ identity matrix.

Then $E(x) = 0$ and $E(xx') = I$.

Let $y$ be defined as

$$y = Ax,$$  \hspace{1cm} (1)

where $A$ is an $N \times N$ matrix.

From Eq. (1) follows $E(y) = E(Ax) = AE(x) = 0$ and

$$\Sigma \equiv E(yy') = E(Axx'A') = AE(xx')A' = AA'.$$

For a known positive definite $N \times N$ matrix $\Sigma$ an $N \times N$ matrix $A$ must be found such that $\Sigma = AA'$. 
a. Monte Carlo - Cholesky Decomposition

Use the *Cholesky decomposition* of $\Sigma$ (in Matlab $\text{chol}(\Sigma)'$)

- $A = \text{chol}(\Sigma)$
- check that $A$ is a *lower triangular* matrix (as Matlab returns an upper triangular matrix, it has to be transposed)
- $A$ has the property of $\Sigma = AA'$

The multivariate return series can be generated by

$$r = y + \mu,$$

where $\mu$ is a vector of constants (here expected returns).

The random vector $r$ satisfies the properties $E(r) = \mu$ and

$$\text{Cov}(r) = E(rr') - E(r)E(r') = E(y + \mu)(y + \mu)' - \mu\mu' = E(yy') + 2\mu E(y') + \mu\mu' - \mu\mu' = E(yy') = \Sigma.$$
The following experiment is based on DeMiguel/Nogales (2007).

Given are the (theoretical) moments of monthly returns

$$\bar{\mu} = \begin{pmatrix} 0.01 \\ 0.01 \\ 0.01 \\ 0.01 \end{pmatrix}, \quad \Sigma = 0.04 \times I_{N \times N}.$$

The Mean-Variance optimal weights for any risk aversion coefficient are 0.25 for each asset!\(^1\)

---

\(^1\)To verify this, set $\bar{\mu} = a\mathbf{1}$ and $\Sigma = bI$ in Eq. (3).
b. Experiment

Generate 140 returns $r_i \sim \mathcal{N}(\bar{\mu}, \Sigma)$, $i = 1 \ldots 140$.

For different values of $t$ calculate sample moments from

$$\hat{\mu} = \frac{1}{120} \sum_{i=t}^{120+t-1} r_i, \quad \hat{\Sigma} = \frac{1}{120-1} \sum_{i=t}^{120+t-1} (r_i - \hat{\mu})(r_i - \hat{\mu})'$$

Rolling window estimation:

Rebalancing date $t = 1$: Calculate $\hat{\mu}$ and $\hat{\Sigma}$ using $r_1 \ldots r_{120}$
Rebalancing date $t = 2$: Calculate $\hat{\mu}$ and $\hat{\Sigma}$ using $r_2 \ldots r_{121}$
... until $t = 20$.

In the following, for each $\hat{\mu}$, $\hat{\Sigma}$ optimal weights are calculated.
\( \bar{\mu} \) and \( \Sigma \) are unknown:

\( \bar{\mu} \) and \( \Sigma \) are estimated using 120 monthly returns \( r_i \) (rolling window). The dashed line shows optimal portfolio weights without estimation error (\( \gamma = 0.25 \)) and \( \gamma = 1 \).
\( \bar{\mu} \) is estimated and \( \Sigma \) is known:
c. Results

\[ \bar{\mu} \text{ is known and } \Sigma \text{ is estimated:} \]

![Graph showing weights over rebalancing dates with T=120](image)
c. Results

Remarks:

- Estimation error can cause strong deviation of estimated portfolio weights from theoretically optimal weights. In the above example where $\bar{\mu}$ and $\Sigma$ are unknown the maximal estimated weight is 0.78 and the minimal weight is $-0.09$ while (theoretically) optimal weights are all 0.25.

- Estimation error in expected returns has larger impact on asset allocation than estimation error in the covariance matrix.

- Higher risk aversion leads to a portfolio allocation closer to the global minimum variance portfolio (GMVP) (see next section). As the GMVP does not rely on expected returns, higher risk aversion reduces estimation error (besides reducing portfolio risk)!
2. Robust Optimization
The previous section has shown the effects of estimation error in expected returns and the covariance matrix on optimal asset allocation.

Optimal weights using sample estimates (esp. of expected returns) can strongly deviate from their theoretical values.

This section shows a robust portfolio optimization method based on Scherer (2007) (and Tütüncü and König (2004)) that mitigates the problem of estimation error.
In traditional mean variance optimization of a portfolio with \( N \) assets, expected utility is maximized using the following Lagrangian function

\[
L = \bar{\mu}'w - \frac{1}{2}\gamma w'\Sigma w - \lambda (w'1 - 1)
\]  

(2)

where

- \( \bar{\mu} \): is an \( N \times 1 \) vector of expected returns,
- \( \Sigma \): is the \( N \times N \) covariance matrix of assets’ returns,
- \( w \): is an \( N \times 1 \) vector of portfolio weights,
- \( \gamma \): is the parameter of relative risk aversion,
- \( \lambda \): is a Lagrange multiplier.

Note that \( (w'1 - 1) \) is the constraint requiring that the optimal weight vector’s elements sum up to 1.
Traditional Optimization

The weight vector that maximizes Eq. (2) can be found by

$$\frac{\partial L}{\partial w} = \bar{\mu} - \gamma \Sigma w - \lambda 1 = 0$$
$$\frac{\partial L}{\partial \lambda} = w' 1 - 1 = 0$$

The solution is given by

$$w_{mv}^* = \frac{1}{\gamma} \Sigma^{-1} \left( \bar{\mu} - \frac{1' \Sigma^{-1} \bar{\mu}}{1' \Sigma^{-1} 1} 1 \right) + \frac{\Sigma^{-1} 1}{1' \Sigma^{-1} 1} w_{\text{spec}}^* + \frac{\Sigma^{-1} 1}{1' \Sigma^{-1} 1} w_{\text{min}}^*$$

(3)

where $w_{\text{min}}^*$ is the portfolio weight of the global minimum variance portfolio and $w_{\text{spec}}^*$ is a speculative demand.

- Estimation error in sample means \( \bar{\mu} \) has larger effect on portfolio allocation than estimation error in the covariance matrix.

- Instead of plugging the mean vector into the Mean-Variance optimizer, we will use a ‘worst case’ vector \( \mu_* \).

- The worst case expected portfolio return \( w' \mu_* \) is lower than (or equal to) \( w' \bar{\mu} \); this makes it generally less attractive to invest into the speculative weight \( w_{\text{spec}}^* \) which depends on \( \bar{\mu} \) (or \( \mu_* \)).

- The robust optimization method presented here leads to higher investment into the global minimum variance portfolio \( w_{\min}^* \) than in traditional Mean-Variance optimization.
Robust Optimization – a. Worst Case Expected Returns

The vector $\mu$ is drawn from a multivariate normal distribution

$$\mu \sim \mathcal{N}(\bar{\mu}, \Omega),$$

where $\Omega$ is an $N \times N$ positive definite covariance matrix.

We consider a realization of $\mu$ as an extreme event if

$$(\mu - \bar{\mu})' \Omega^{-1} (\mu - \bar{\mu}) \geq \kappa^2_{1-\alpha, N}.$$

Note that $(\mu - \bar{\mu})' \Omega^{-1} (\mu - \bar{\mu})$ is $\chi^2_N$ distributed (chi-square distributed with $N$ degrees of freedom – $N$ is the number of assets).
where \( \kappa_{1-\alpha,N}^2 = \text{Inv-} \chi^2_N \left( 1 - \alpha \right) \) and \( \text{Inv-} \chi^2_N \) is the inverse chi-square cumulative density function with \( N \) degrees of freedom.

Example: \( \kappa_{0.95,4}^2 = 9.4877 \). (use \( \text{chi2inv}(0.95,4) \) in Matlab).
Robust Optimization – a. Worst Case Expected Returns

Now consider the following set of expected return vectors

$$((\mu - \bar{\mu})' \Omega^{-1} (\mu - \bar{\mu})) = \kappa_{1-\alpha,N}^2.$$  (4)

Geometrically, Eq. (4) describes an $N$-dimensional ellipsoid. We will restrict our analysis to expected return vectors $\mu$ that lie on this ellipsoid.

On this ellipsoid the value of $\mu$ with the lowest expected portfolio return is obtained by maximizing the Lagrangian function

$$L(\mu, \lambda) = w' \bar{\mu} - w' \mu - \frac{\lambda}{2} \left( (\mu - \bar{\mu})' \Omega^{-1} (\mu - \bar{\mu}) - \kappa_{1-\alpha,N}^2 \right)$$

where $w$ is any portfolio weight vector.
Robust Optimization – a. Worst Case Expected Returns

Setting $\partial L / \partial \mu = 0$ and $\partial L / \partial \lambda = 0$ and solving this system of linear equations with respect to $\mu$ yields

$$
\mu_\ast = \bar{\mu} - \frac{\kappa_{1-\alpha, N}}{\sqrt{w'\Omega w}} \Omega w
$$

and

$$
w'\mu_\ast = w'\bar{\mu} - \kappa_{1-\alpha, N} \sqrt{w'\Omega w}.
$$

From Eq. (5) it is obvious that $w'\mu_\ast \leq w'\bar{\mu}$.

The value $w'\mu_\ast$ is particularly low for high values of $\kappa_{1-\alpha, N}$.

For $\Omega = \frac{1}{T} \Sigma$ Eq. (5) can be rewritten as

$$
w'\mu_\ast = w'\bar{\mu} - \kappa_{1-\alpha, N} T^{-1/2} \sqrt{w'\Sigma w},
$$

where $\Sigma$ is the covariance matrix of assets’ returns.
The traditional Mean-Variance objective is to maximize the Lagrangian

\[ L = \bar{\mu}'w - \frac{1}{2}\gamma w'\Sigma w - \lambda(w'1 - 1). \]

\[ L_{rob} = \mu_* w - \frac{1}{2}\gamma w'\Sigma w - \lambda(w'1 - 1) \]

\[ = \bar{\mu}'w - \kappa_{1-\alpha,N}T^{-1/2}\sqrt{w'\Sigma w} - \frac{1}{2}\gamma w'\Sigma w - \lambda(w'1 - 1). \]
Robust Optimization – b. Deriving Optimal Weights

Overview
1. Estimation Error
2. Robust Optimization
   a. Worst Case \( \mu \)
   b. Derivation
   c. Example
3. Reverse Optimization

Solving \( \partial L_{rob}/\partial w = 0 \) and \( \partial L_{rob}/\partial \lambda = 0 \) gives

\[
\mathbf{w}_{rob}^* = \left( 1 - \frac{T^{-1/2} \kappa_{1-\alpha, N}}{\gamma \sigma^*_p + T^{-1/2} \kappa_{1-\alpha, N}} \right) \frac{1}{\gamma} \Sigma^{-1} \left( \bar{\mu} - \frac{1'}{\Sigma^{-1}} \frac{\mu}{1'} \frac{1}{1} \right)
\]

\[
+ \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}' \Sigma^{-1} \mathbf{1}}
\]

\[
\mathbf{w}_{rob}^* = \left( 1 - \frac{T^{-1/2} \kappa_{1-\alpha, N}}{\gamma \sigma^*_p + T^{-1/2} \kappa_{1-\alpha, N}} \right) \mathbf{w}^*_{spec} + \mathbf{w}^*_{min} \quad (7)
\]

Note that \( \sigma^*_p = \sqrt{\mathbf{w}_{rob}^{*'} \Sigma \mathbf{w}_{rob}^*} \) which makes Eq. (7) not directly solvable. However, Eq. (7) can be solved numerically.
Moreover, note that

- If $\kappa_{1-\alpha,N} \to 0$ or $T \to \infty$,
  \[
  w_{\text{rob}}^* = \frac{1}{\gamma} \Sigma^{-1} \left( \bar{\mu} - \frac{1'}{1'} \Sigma^{-1} \bar{\mu} 1 \right) + \frac{\Sigma^{-1} 1}{1' \Sigma^{-1} 1} = w_{\text{mv}}^*
  \]

- If $\kappa_{1-\alpha,N} \to \infty$ or $T \to 0$,
  \[
  w_{\text{rob}}^* = \frac{\Sigma^{-1} 1}{1' \Sigma^{-1} 1} = w_{\text{min}}^*
  \]
Impact of estimation error in traditional M-V Optimization

\[ (\gamma = 1) \]
Impact of estimation error in robust M-V Optimization

\[ T=120 \]

\[
\begin{array}{c}
\text{weights} \\
\end{array}
\]

\[
\begin{array}{c}
\text{rebalancing dates} \\
\end{array}
\]

\[ (\gamma = 1) \]
Robust Optimization – c. Example

Investment into $w^*_{\text{spec}}$ increases with higher sample size $T$.
(Note again that $\sigma^*_p$ is not constant in Eq. (7) - the curve below was simulated numerically.)

\[
\psi = \left( 1 - \frac{T^{-1/2} \kappa_{1-\alpha,N}}{\gamma \sigma^*_p + T^{-1/2} \kappa_{1-\alpha,N}} \right) \leq 1
\]
Robust Optimization – c. Example

Remarks:

□ as $w^\prime \mu_* \leq w^\prime \bar{\mu}$, robust optimization using $\mu_*$ leads to ‘less attractive’ expected returns which reduces speculative demand $w^*_{\text{spec}}$. This is shown by Eq. (7), where $\psi \leq 1$ (compare also Eq. (8)).

□ Thus, the fraction of wealth invested into the global minimum variance portfolio $w^*_{\text{min}}$ increases with robust optimization.

□ As $w^*_{\text{min}}$ does not depend on $\bar{\mu}$, the resulting portfolio allocation is more robust than in traditional Mean-Variance optimization.

□ This type of robust optimization is equivalent to traditional optimization where the risk aversion coefficient is increased to $\frac{1}{\psi} \gamma$. Note that $\psi$ depends on the ‘uncertainty aversion’ $\kappa_{1-\alpha,N}$ and $T$. 
3. Reverse Optimization
Reverse optimization (applied in e.g. Black and Litterman (1992)) avoids the problem arising from estimation error in expected returns.

Traditional Mean-Variance optimization uses expected returns (and the covariance matrix) as input and provides portfolio weights as output.

In contrast, reverse optimization uses portfolio weights (and the covariance matrix) as input and provides expected returns as output.

The rationale behind using portfolio weights as input is that in some cases weights are easier to estimate than expected returns.
Reverse Optimization

As in Black/Litterman (1992) we confine our analysis to expected excess returns $\mu_e$ (expected returns minus the riskfree rate).

Then in Eq. (2) the constraint $\lambda(w'1 - 1)$ can be dropped as weights do not have to sum up to 1 (because implicitly, remaining wealth is invested into the riskfree asset)

$$L = \mu'_e w - \frac{1}{2} \gamma w' \Sigma w$$

This simplifies the solution of the optimal weight. Setting $\partial L/\partial w = 0$ gives

$$\mu_e - \gamma \Sigma w = 0.$$ 

Rearranging yields

$$w^* = \frac{1}{\gamma} \Sigma^{-1} \mu_e.$$

(9)
Reverse Optimization

From Eq. (9) follows the main equation for reverse optimization:

$$\mu_e^* = \gamma \Sigma w^*.$$

In e.g. the Black/Litterman model value weighted market portfolio weights (or strategic asset allocation weights) are used for $w^*$. 
Remarks:

- By definition, plugging $\mu_e^*$ into Eq. (9) yields $w^*$.

- However, when changing the value of the risk aversion coefficient $\gamma$, optimal weights obtained in Eq. (9) differ from the original reference weights $w^*$.

- A ‘sensible’ value should be chosen for the reference weights $w^*$. This improves the economic intuition of the results.
Advantages:

- The advantage of reverse optimization is that it does not rely on the estimation of expected returns. This leads to portfolio allocations more robust than in traditional Mean-Variance optimization.

- Resulting weights are often economically easier to interpret than weights from traditional Mean-Variance optimization (for reasonable reference weights $w^*$).

Disadvantage:

- Reverse optimization may throw away useful information contained in (estimated) expected returns.

