Robust portfolio optimization under multiperiod mean-standard deviation criterion

Spiridon Penev\textsuperscript{1} Pavel Shevchenko\textsuperscript{2} Wei Wu\textsuperscript{1}

\textsuperscript{1}The University of New South Wales, Sydney
Australia

\textsuperscript{2}Macquarie University
Sydney, Australia

Wollongong, July 2017
Outline

1. Introduction
2. Motivation
   - Technicalities
3. Main result
   - Quantification of Model Risk with Empirical Data.
5. Conclusion.
6. References
We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.

Uncertainty is measured by the Kullback-Leibler divergence (more generally by $\alpha$-divergence).

We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.

As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.

Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution.
We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.

Uncertainty is measured by the Kullback-Leibler divergence (more generally by $\alpha$-divergence).

We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.

As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.

Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution.
We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.

Uncertainty is measured by the Kullback-Leibler divergence (more generally by $\alpha$-divergence).

We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.

As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.

Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution.
We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.

Uncertainty is measured by the Kullback-Leibler divergence (more generally by $\alpha$-divergence).

We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.

As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.

Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution → quantify model risk.
We look into model risk, defined as the loss due to uncertainty of the underlying distribution of the return of the assets in a portfolio.

Uncertainty is measured by the Kullback-Leibler divergence (more generally by $\alpha$-divergence).

We show that in the worst case scenario, the optimal robust strategy can be obtained in a semi-analytical form.

As a consequence, we quantify model risk. By combining with a Monte Carlo approach, the optimal robust strategy can be calculated numerically.

Numerically compare performance of the robust strategy with the optimal non-robust strategy, the latter being calculated at a nominal distribution → quantify model risk.
Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): \( \max_u (E(W) - \kappa \cdot \text{Var}(W)) \)
- Others: Safety-first: \( \min_u P(W \leq d) \), targeting particular wealth level: \( \max_u E(W - \hat{W}^2) \), etc.
- One more: Mean-st. deviation (MSD) criterion: \( \max_u (E(W) - \kappa \sqrt{\text{Var}(W)}) \).

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizing the MSD criterion (with a suitable \( \kappa \)).
Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): $\max_u (E(W) - \kappa \cdot Var(W))$
- Others: Safety-first: $\min_u P(W \leq d)$, targeting particular wealth level: $\max_u E(W - \hat{W}^2)$, etc.
- One more: Mean-st. deviation (MSD) criterion: $\max_u (E(W) - \kappa \sqrt{Var(W)})$.

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizing the MSD criterion (with a suitable $\kappa$).
Which criterion should an investor choose to optimize portfolio wealth?

- Markowitz (1952): \( \max_u (E(W) - \kappa \cdot Var(W)) \)
- Others: Safety-first: \( \min_u P(W \leq d) \), targeting particular wealth level: \( \max_u E(W - \hat{W}^2) \), etc.
- One more: Mean-st. deviation (MSD) criterion: \( \max_u (E(W) - \kappa \sqrt{Var(W)}) \).

Why the latter: for elliptically distributed returns, optimizing a risk measure form the whole class of Translation-invariant and positive-homogeneous risk measures (TIPH) is equivalent to optimizing the MSD criterion (with a suitable \( \kappa \)).
However: i) the joint distribution of the assets is unknown (e.g., ‘slightly deviating’ from a nominal multivariate normal). ii) changes of the distribution over time (in multiple periods) (need dynamic approach).

The deviation can be measured by: KL divergence, $\alpha$-divergence..

Try to quantify the intuition: “big divergence $\rightarrow$ significant impact on an optimal investment decision that is calculated under the nominal distribution”.

Ultimate goal: If distributional assumptions are violated only ‘slightly’ $\rightarrow$ use the optimal investment strategy under the nominal model” (since robust approach may deliver too pessimistic strategy). Ideally: ball of radius $\eta_0$ around the nominal model: stay with the nominal “inside”, switch to robust “outside”. Need to quantify model risk from risk management perspective.
Outline

1 Introduction

2 Motivation
   ■ Technicalities

3 Main result

4 Numerical Examples.
   ■ Quantification of Model Risk with Empirical Data.

5 Conclusion.

6 References
Using KL divergence: reasonable for short term re-balancing (daily or weekly). Big advantage: closed form worst case distribution is available→allows us to get the optimal (time-consistent) strategy under the worst case distribution in a semi-analytical form. Our previous work (BGPW, *Automatica* 2016) delivers the optimal strategy under the nominal model→can compare performance under worst case scenario.
Problem Formulation - $d > 1$ risky assets; fixed investment horizon $[0, N]$; return of each asset over the $n$th period $[n, n+1], n = 0, ..., N - 1$: as $r_{n+1} = (r_{n+1}^1, ..., r_{n+1}^d)^T$, with $r_{n+1}^i, i = 1, ..., d$. Filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ with the sample space $\Omega$, the sigma-algebra $\mathcal{F}$, filtration $(\mathcal{F}_n)$, probab. measure $\mathbb{P}$, sigma-algebra $\mathcal{F}_n = \sigma(r_m, 1 \leq m \leq n)$. Return vector $r_{n+1}: \mathbb{E}(|r_{n+1}|^2) < \infty$.

At time $n = 0, ..., N - 1$: re-balance using strategy $u = (u_0, ..., u_{N-1})^T$, all $u_n$, taking values in $U$:

$$U = \left\{ u \in \mathbb{R}^d : 1^T u = 1, \text{ for } i = 1, ..., d \right\}.$$

For $m > 0$, we use $\mathcal{U}^m$ to denote the set of admissible sub-strategies $u^m = (u_n)_{n \geq m}$.
Let $W_n$ denote the wealth at time $n$ ($n = 0, ..., N$). Assume: $W_n$ and $r_{n+1}$ are independent. During $[n, n+1]$, wealth changes:

$$W_{n+1} = W_n (1 + r_{n+1})^T u_n = W_n R_{n+1}^T u_n,$$

where $R_{n+1} = 1 + r_{n+1}$. At any time $m$, aim: optimize $J_{m,x}(u^m) = \mathbb{E}\left( \sum_{n=m}^{N-2} \mathcal{J}_{n,w}(W_{n+1}) + \mathcal{J}_{N-1,w_{N-1}}(W_N) | W_m = x \right)$, where

$$\mathcal{J}_{n,w}(W_{n+1}) = W_{n+1} - \kappa_n \sqrt{\text{Var}_{n,w}(W_{n+1})}$$

$$\text{Var}_{n,w}(W_{n+1}) = \text{Var}(W_{n+1} | W_n), \quad \text{and} \quad \Sigma_n = \text{Var}(r_{n+1}),$$

The above: multi-period selection criterion of MSD type; $\kappa_n$ characterizes investor’s risk aversion. Details: BGPW.
The value function of this control problem:

$$V(m, x) = \sup_{u^m \in U^m} J_{m, x}(u^m).$$  \tag{1}$$

Use KL divergence $\mathcal{R}(\mathcal{E}) = \mathbb{E} \left( \mathcal{E} \log \mathcal{E} \right)$, where $\mathcal{E}$ is the ratio of the density of an alternative distribution to model distribution. For a given $\eta > 0$, a KL divergence ball is:

$$B_\eta = \{ \mathcal{E} : \mathcal{R}(\mathcal{E}) \leq \eta \}.$$  \tag{2}$$
Next: robust version of the problem. Let balls

$$\mathcal{B}_{\eta_n} = \{ \mathcal{E} : \mathcal{R}(\mathcal{E}) \leq \eta_n \}, \text{ where } n = 0, \ldots, N - 1.$$

Given any starting time $m = 0, \ldots, N - 1$, we denote the set of $\mathcal{E}^m = (\mathcal{E}_m, \ldots, \mathcal{E}_{N-1})$ such that each $\mathcal{E}_n \in \mathcal{B}_{\eta_n}$, where $n = m, \ldots, N - 1$, by $\mathcal{B}^m (\mathcal{E}_n : \text{ratio of density of alternative to density of the nominal distribution over } [n, n+1])$. Then:

$$V(m, x) = \sup_{u^m \in \mathcal{U}^m} \inf_{\mathcal{E}^m \in \mathcal{B}^m} J_{m,x}(\mathcal{E}^m, u^m),$$

$$J_{m,x}(\mathcal{E}^m, u^m) = \mathbb{E} \left( \mathcal{E}_m \mathcal{W}_m \left( R_{m+1}^T u_m \right) - \kappa_m \sqrt{u_m^T \Sigma_m u_m} \right) + \sum_{n=m+1}^{N-1} e^{-\eta_n \gamma_n} \mathcal{E}_n \mathcal{I}_n, \mathcal{W}_n(W_{n+1} | \mathcal{W}_m = x).$$
Remark. $\gamma_n$ is a weighting to reflect the fact that a large uncertainty in the future should not impact too much the decision at the current stage. In addition, this also guarantees the existence of an optimal solution. As $\eta_n \to 0 \to$ return to the non-robust case.
Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

Definition

A strategy $u_{m,*} = (u_{m,*},...,u_{N-1,*})$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(E^m, u^m)$ if:

1: Let $\mathcal{A}^m \subset \mathcal{U}^m$ be a set of strategies $u^m = (v, u_{m+1,*},...,u_{N-1,*})$. Then $\exists E^m,* \in \mathcal{B}^m$ s.t. $\sup_{u^m \in \mathcal{A}^m} \inf_{E^m \in \mathcal{B}^m} J_{m,x}(E^m, u^m) = J_{m,x}(E^m,*, u^m,*).$

2: For $n = m+1,...,N-1$, $\exists E^n,* \in \mathcal{B}^n$ s.t. $\sup_{u^n \in \mathcal{U}^n} \inf_{E^n \in \mathcal{B}^n} J_{n,x}(E^n, u^n) = J_{n,x}(E^n,*, u^n,*).$

If only 1 is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$. 
Main result

Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

Definition

A strategy $u_{m,*} = (u_{m,*}, ..., u_{N-1,*})$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(E^m, u^m)$ if:

1: Let $\mathcal{A}^m \subset \mathcal{U}^m$ be a set of strategies $u^m = (v, u_{m+1,*}, ..., u_{N-1,*})$. Then $\exists E_{m,*} \in \mathcal{B}^m$ s.t.
$$\sup_{u^m \in \mathcal{A}^m} \inf_{E^m \in \mathcal{B}^m} J_{m,x}(E^m, u^m) = J_{m,x}(E_{m,*}, u_{m,*}).$$

2: For $n = m + 1, ..., N - 1$, $\exists E_{n,*} \in \mathcal{B}^n$ s.t.
$$\sup_{u^n \in \mathcal{U}^n} \inf_{E^n \in \mathcal{B}^n} J_{n,x}(E^n, u^n) = J_{n,x}(E_{n,*}, u_{n,*}).$$

If only 1 is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$.
Semi-Analytical Optimal Solution under KL Divergence. We require strongly time consistent optimal robust strategy. It represents a robustified version of a strong time consistent optimal strategy inspired by Kang and Filar 2006 (see also BGPW).

**Definition**

A strategy $u^m, \ast = (u_m^*, ..., u_{N-1}^*)$ is strongly time consistent optimal robust w.r.t. $J_{m,x}(\mathcal{E}^m, u^m)$ if:

1. Let $\mathcal{A}^m \subset \mathcal{U}^m$ be a set of strategies $u^m = (v, u_{m+1}^*, ..., u_{N-1}^*)$. Then $\exists \mathcal{E}^m, \ast \in \mathcal{B}^m$ s.t.
   \[
   \sup_{u^m \in \mathcal{A}^m} \inf_{\mathcal{E}^m \in \mathcal{B}^m} J_{m,x}(\mathcal{E}^m, u^m) = J_{m,x}(\mathcal{E}^m, \ast, u^m, \ast).
   \]

2. For $n = m + 1, ..., N - 1$, $\exists \mathcal{E}^n, \ast \in \mathcal{B}^n$ s.t.
   \[
   \sup_{u^n \in \mathcal{U}^n} \inf_{\mathcal{E}^n \in \mathcal{B}^n} J_{n,x}(\mathcal{E}^n, u^n) = J_{n,x}(\mathcal{E}^n, \ast, u^n, \ast).
   \]

If only 1 is satisfied: call it weakly time consistent optimal robust strategy w.r.t. $J_{n,x}(\cdot)$. 
Since the value function of the robust control problem is separable (in the sense that it can be written as a sum of expectations), we know: weakly time consistent optimal strategy, which can be found by period-wise optimization, is also a strongly time consistent optimal strategy.
Theorem. Suppose \((u_m^*, m = 0, 1, \ldots, N - 1)\) is a strategy where there exists a sequence \((\theta_m^*)\) s.t. \(\mathbb{E}\left(\exp\left(-R_{m+1}^T u_m^* \frac{2}{\theta_m^*}\right)\right) < \infty\), and

\[
\begin{align*}
    u_m^* &= \frac{S_m^*}{\kappa_m} \left(\Sigma_m^{-1} X_m^* - \frac{b_m^* \Sigma_m^{-1} 1}{a_m}\right) + \frac{\Sigma_m^{-1} 1}{a_m}, \\
    S_m^* &= \sqrt{\frac{1}{a_m}} \left(1 - \frac{h_m^*}{\kappa_m^2} + \frac{(b_m^*)^2}{\kappa_m^2 a_m}\right) = \sqrt{\frac{1}{a_m}} \left(1 - \frac{1}{\kappa_m^2} g_m^*\right), \\
    X_m^* &= \frac{\mathbb{E}\left(\exp\left(-R_{m+1}^T u_m^* \frac{1}{\theta_m^*}\right) R_{m+1}\right)}{\mathbb{E}\left(\exp\left(-R_{m+1}^T u_m^* \frac{1}{\theta_m^*}\right)\right)} + e^{-\eta_{m+1, m+1}} G_{m+1}(u_{m+1}^*, \theta_{m+1}^*) \mathbb{E}(R_{m+1}),
\end{align*}
\]

\(\mathbb{E}(\mathcal{E}_m^* \log(\mathcal{E}_m^*)) = \eta_m\),
where

\[
g_m^* = h_m^* - \frac{(b_m^*)^2}{a_m}, \quad h_m^* = (X_m^*)^T \Sigma_m^{-1} X_m^*, \quad a_m = 1^T \Sigma_m^{-1} 1,
\]

\[
b_m^* = 1^T \Sigma_m^{-1} X_m^*, \quad \mathcal{E}_m^* = \frac{\exp \left( - R_{m+1}^T u_m^* \frac{1}{\theta_m^*} \right)}{\mathbb{E} \left( \exp \left( - R_{m+1}^T u_m^* \frac{1}{\theta_m^*} \right) \right)} \mathbb{P}\text{-a.s.},
\]

\[
G_m(u_m^*, \theta_m^*) = -\theta_m^* \log \mathbb{E} \left( \exp \left( - R_{m+1}^T u_m^* \frac{1}{\theta_m^*} \right) \right) + e^{-\eta_m} \gamma_{m+1} G_{m+1}(u_{m+1}^*, \theta_{m+1}^*) \mathbb{E} \left( R_{m+1}^T u_m^* \right) - \kappa_m S_m - \eta_m \theta_m^*,
\]

\[
G_N(u_N^*, \theta_N^*) = 0.
\]

Then, \((u_m^*)\) is optimal, and the value function is given by

\[
V(m, x) = x G_m(u_m^*, \theta_m^*),
\]

where \(x \in (0, \infty)\).
Numerical Examples.

$d = 3$ stocks: Navitas, Domino and Tabcorp, $N = 5$. Historical
daily prices collected: 1 Jan 2015 - 31 Dec 2015. The 261 daily
returns calculated. Set $\kappa_n = 3$, $W_0 = 1$, returns are assumed to
be i.i.d. over the investment horizon.

Comparison of Optimal Robust and Non-Robust Portfolio. For a
nominal 3-dim. MVN, mean $\mu$, cov. mat. $\Sigma$, and alternative
model: 3-dimensional MVN, mean $\bar{\mu}$, cov. mat. $\bar{\Sigma}$, KL:

$$R(E) = \frac{1}{2} \left( \text{trace}(\Sigma^{-1}\bar{\Sigma}) + (\mu - \bar{\mu})^T \Sigma^{-1} (\mu - \bar{\mu}) - d + \log \left( \frac{\det(\Sigma)}{\det(\bar{\Sigma})} \right) \right).$$

For illustration→ consider
$\bar{\mu} = c_n \times \mu$, for some $c_n \in \mathbb{R}$, and $\bar{\Sigma} = \Sigma$. Assume expected
returns and covariance matrices stay constant over the
investment horizon (i.e. $\mathbb{E}(r_n) = \mu$, and $\Sigma_n = \Sigma$).
In the worst case scenario for model disturbance, the alternative model is on the boundary of the KL div. ball (see Theorem) → the divergence between the two models is equal to $\eta_n$ → we can find $c_n$. Assume $\eta_n$ to be constant over the entire investment horizon. For simplicity, simply write $\eta$ and $c$. Choose $(\gamma_2, \gamma_3, \gamma_4, \gamma_5)$ such that $(\eta \gamma_2, \eta \gamma_3, \eta \gamma_4, \eta \gamma_5) = (7.5, 8.0, 8.5, 9.0)$ (reflecting investor’s risk tolerance for uncertainty of distribution) (in contrast to $\kappa$ which is the risk aversion of the investor’s preference for a fixed distribution). So: the investor will have its own freedom to choose the amount of penalization.
Suppose: the alternative model is the true model. By generating data from the alternative model, compare the performance under the optimal robust and non-robust strategies for different \( \eta \). The optimal robust strategy: calculated by using 500,000 Monte Carlo simulations. Then, we simulate 500,000 daily return paths (over 5 days), and compare the performance over three reasonable metrics:
Numerical Examples.

Figure: the number of times robust outperforms non-robust
Numerical Examples.

Figure: robust vs non-robust: expected terminal wealth
Numerical Examples.

**Figure:** robust vs non-robust: ratio of the difference between the expected terminal wealth and the initial wealth to the standard deviation of the terminal wealth.
Numerical Examples.

Table: Performance of Robust and Non-Robust Optimal Solution (Shift of Mean)

<table>
<thead>
<tr>
<th>$c$</th>
<th>$\eta$</th>
<th>Times robust outperforms model</th>
<th>%</th>
<th>$E(W_N)$ robust</th>
<th>$E(W_N)$ non-robust</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2139</td>
<td>0.0050</td>
<td>244429</td>
<td>48.89%</td>
<td>1.0015</td>
<td>1.0016</td>
<td>-0.0001</td>
</tr>
<tr>
<td>-1.4859</td>
<td>0.0500</td>
<td>285828</td>
<td>57.17%</td>
<td>0.9903</td>
<td>0.9891</td>
<td>0.0012</td>
</tr>
<tr>
<td>-2.5156</td>
<td>0.1000</td>
<td>309814</td>
<td>61.96%</td>
<td>0.9842</td>
<td>0.9816</td>
<td>0.0026</td>
</tr>
<tr>
<td>-3.9718</td>
<td>0.2000</td>
<td>336583</td>
<td>67.32%</td>
<td>0.9761</td>
<td>0.9711</td>
<td>0.0050</td>
</tr>
<tr>
<td>-5.0892</td>
<td>0.3000</td>
<td>362909</td>
<td>72.58%</td>
<td>0.9715</td>
<td>0.9631</td>
<td>0.0084</td>
</tr>
<tr>
<td>-6.0312</td>
<td>0.4000</td>
<td>378952</td>
<td>75.79%</td>
<td>0.9678</td>
<td>0.9564</td>
<td>0.0114</td>
</tr>
<tr>
<td>-6.8611</td>
<td>0.5000</td>
<td>391459</td>
<td>78.29%</td>
<td>0.9650</td>
<td>0.9505</td>
<td>0.0145</td>
</tr>
</tbody>
</table>
Another case with a closed form formula for the KL div.: when both the nominal and alternative models are multivariate skew-normal. Given a nominal model $Y \sim SN_d(\mu, \Sigma, \xi)$ and an alternative model $\tilde{Y} \sim SN_d(\tilde{\mu}, \tilde{\Sigma}, \tilde{\xi})$, then models is given by:

$$R_{skew}(E) = R(E) + 2\sqrt{\frac{2}{\pi}}(\mu - \tilde{\mu})^T \Sigma^{-1} \Sigma^{\frac{1}{2}} \tilde{\xi} - \mathbb{E} \left( \log \left( 2\Phi(\Xi_2 | 1 - \xi^T \xi) \right) \right)$$

$$+ \mathbb{E} \left( \log \left( 2\Phi(\Xi_1 | 1 - \tilde{\xi}^T \tilde{\xi}) \right) \right),$$

$\Xi_1 \sim SN_1(0, \tilde{\xi}^T \tilde{\xi}, \sqrt{\tilde{\xi}^T \tilde{\xi}})$,

$\Xi_2 \sim SN_1 \left( \xi^T \Sigma^{-\frac{1}{2}} (\mu - \tilde{\mu}), \xi^T \Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{1}{2}} \xi, \frac{\xi^T \Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{1}{2}} \xi}{\sqrt{\xi^T \Sigma^{-\frac{1}{2}} \Sigma \Sigma^{-\frac{1}{2}} \xi}} \right)$. 
Numerical Examples.

Figure: the number of times robust outperform non-robust
Figure: robust vs non-robust: expected terminal wealth
Figure: robust vs non-robust: ratio of the difference between the expected terminal wealth and the initial wealth to the standard deviation of the terminal wealth.
**Table:** Performance of Robust and Non-Robust Optimal Solution (skew-normal)

<table>
<thead>
<tr>
<th>$\tilde{c}$</th>
<th>$\eta$</th>
<th>Times robust outperforms model</th>
<th>%</th>
<th>$E(W_N)$ robust</th>
<th>$E(W_N)$ non-robust</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.7299</td>
<td>0.0050</td>
<td>302353</td>
<td>60.88%</td>
<td>0.9643</td>
<td>0.9635</td>
<td>0.0008</td>
</tr>
<tr>
<td>-0.7125</td>
<td>0.0500</td>
<td>300014</td>
<td>60.00%</td>
<td>0.9659</td>
<td>0.9644</td>
<td>0.0015</td>
</tr>
<tr>
<td>-0.6929</td>
<td>0.1000</td>
<td>297997</td>
<td>59.60%</td>
<td>0.9674</td>
<td>0.9653</td>
<td>0.0021</td>
</tr>
<tr>
<td>-0.6524</td>
<td>0.2000</td>
<td>291802</td>
<td>58.36%</td>
<td>0.9697</td>
<td>0.9673</td>
<td>0.0023</td>
</tr>
<tr>
<td>-0.6097</td>
<td>0.3000</td>
<td>290593</td>
<td>58.12%</td>
<td>0.9724</td>
<td>0.9694</td>
<td>0.0030</td>
</tr>
<tr>
<td>-0.5643</td>
<td>0.4000</td>
<td>286906</td>
<td>57.38%</td>
<td>0.9748</td>
<td>0.9717</td>
<td>0.0031</td>
</tr>
<tr>
<td>-0.5155</td>
<td>0.5000</td>
<td>282951</td>
<td>56.59%</td>
<td>0.9773</td>
<td>0.9741</td>
<td>0.0032</td>
</tr>
</tbody>
</table>

W.r. to the first performance comparison, we notice that now the number of times that the optimal non-robust strategy outperforms the robust strategy decreases as the radius of the divergence increases. However, the percentages of outperform times are all above 50%.
Outline

1. Introduction
2. Motivation
   - Technicalities
3. Main result
   - Quantification of Model Risk with Empirical Data.
5. Conclusion.
6. References
Although in some cases it may be worth choosing the non-robust optimal strategy, we still need to quantify the amount of model risk involved in doing so. Let $\mathbb{Q}$ denote the probability measure of an alternative model (i.e. the empirical measure for the true distribution). The optimal portfolio is said to have a model risk of $\theta$ with a confidence level $q$ if

$$\mathbb{Q}\left(W_{N}^{\text{non-robust}} - W_{N}^{\text{robust}} \leq -\theta \right) = 1 - q.$$  

(In other words, model risk is the $(1 - q)$th-quantile of the distribution of the difference between the terminal wealth under the non-robust strategy and the robust strategy.)
Figure: The distribution of $(W_N^{\text{non-robust}} - W_N^{\text{robust}})$
Numerical Examples.
Quantification of Model Risk with Empirical Data.

Divide the dataset: the first (dataset 1), is used to estimate the expected value and the covariance matrix of the nominal, which yields:

\[
\tilde{\mu} = \begin{pmatrix} 0.0009 \\ 0.0018 \\ 0.0014 \end{pmatrix}, \quad \tilde{\Sigma} = \begin{pmatrix} 0.0004 & 0.0001 & 0.0001 \\ 0.0001 & 0.0004 & 0.0001 \\ 0.0001 & 0.0001 & 0.0003 \end{pmatrix}
\] (5)

For illustration → assume that the nominal is a 3-dimensional MVN with above mean vector and cov. matrix. Dataset 2 (60 data points) is used to evaluate a forecasted distribution of the incoming daily returns (our estimated alternative model to be used in the following 5 days). Use estimation procedure based on the \(k\)th-nearest-neighbor approach. Each time, a sample of 60 is generated from the MVN and the divergence between MNV and alternative model is estimated by using this sample, and dataset 2. Repeat 100,000 times.
Estimated divergence using $k$th-nearest-neighbor approach:
\[ \hat{R}(\mathcal{E}) = \frac{d}{S} \sum_{i=1}^{S} \log \left( \frac{S y_k(i)}{(S-1)\tilde{y}_k(i)} \right). \]
where $\tilde{y}_k(i)$ is the Euclidean distance of the $k$th-nearest-neighbor of $\tilde{Y}_i$ in $(\tilde{Y}_j)_{j \neq i}$, $y_k(i)$ is the Euclidean distance of the $k$th-nearest-neighbor of $\tilde{Y}_i$ in $(Y_i)$, $d = 3$. Get:
\[ \hat{R}(\mathcal{E}) \approx 0.4337. \]

By knowing the KL divergence, we use a bootstrapping to sample 100,000 data points from dataset 2 to construct the distribution of $W_N^{non-robust} - W_N^{robust}$. The estimated model risk at $q = 95\%$ confidence level is 0.0066, i.e., if the optimal non-robust strategy is applied but the optimal robust strategy turns to be more appropriate, then 95\% of the time we would lose no more than 0.66 cents for every one dollar.
Derived a semi-analytical form of optimal robust strategy for an investment portfolio when uncertainty of distribution of the returns is involved. We have applied our approach to several numerical examples and have suggested whether or not we should take the optimal robust or non-robust strategy. We also define model risk from risk management perspective and present an algorithm for quantifying the model risk by using empirical data. This produces a convenient way of quantifying model risk in practice.
References


