Robust optimization by constructing near-optimal portfolios

Martin van der Schans and Tanita de Graaf

Methodological working paper 2017-02
October 16, 2017
Robust optimization by constructing near-optimal portfolios

Martin van der Schans\textsuperscript{1,*} and Tanita de Graaf\textsuperscript{2,*}

Methodological working paper 2017-02

October 16, 2017

Ortec Finance Research Center
P.O. Box 4074, 3006 AB Rotterdam, The Netherlands,
www.ortec-finance.com

SUMMARY

Many investors use optimization to determine their optimal investment portfolio. Unfortunately, optimal portfolios are sensitive to changing input parameters, i.e., they are not robust. Traditional robust optimization approaches aim for an optimal and robust portfolio which, ideally, is the final investment decision. In practice, however, portfolio optimization supports but seldomly replaces the investment decision process. In this paper, we present an approach that both solves the robustness problem and aims to support rather than replace the investment decision process. The method determines a region with near-optimal portfolios which, especially in light of the robustness problem, are all good allocation decisions. Then, as is already common practice, an investor can bring in expert opinion or additional information to select a preferred near-optimal portfolio. We will show that the region of near-optimal portfolios is significantly more robust than the optimal portfolio itself.

JEL Classification: C61, G11

\textsuperscript{1}Ortec Finance Research Center.
\textsuperscript{2}ORTEC B.V.
\textsuperscript{*}Copyright © 2017 Ortec Finance bv. All rights reserved. No part of this paper may be reproduced, in any form or by any means, without permission from the authors. Short sections may be quoted without explicit permission provided that full credit is given to the source. The views expressed are those of the individual author(s) and do not necessarily reflect the views of Ortec Finance bv.
Robust optimization

Many investors use portfolio optimization to determine their optimal portfolio. Unfortunately, optimal portfolios are sensitive to small changes in the optimization’s input parameters. Although most widely studied for mean variance optimization, where optimal portfolios are very sensitive to the estimated mean and covariance matrix (Frankfurter, Phillips, and Seagle 1971; Michaud 1989; Chopra and Ziemba 1993), sensitivity is a generic problem in portfolio optimization (Kondor, Pafka, and Nagy 2007; Ciliberti, Kondor, and Mézard 2007). As discussed in Hurley and Brimberg (2015), the sensitivity is caused by an interaction of an estimation error in the input and the optimization objective.

The literature proposes several robust optimization approaches to deal with the sensitivity problem. Approaches such as shrinkage (Ledoit and Wolf 2004), robust statistics (Reyna et al. 2005), Black-Litterman inverse optimization (Bertsimas, Gupta, and Paschalidis 2012) and Bayesian optimization (Schöttle, Werner, and Zagst 2010) reduce the sensitivity by reducing variation in the input. Other methods, such as regularization, change the optimization objective to make it less sensitive to varying input (DeMiguel et al. 2009; Brodie et al. 2009). Also, there are hybrid methods that both reduce variation in the input and make the optimization objective less sensitive to varying input. For example, the optimization community proposes a general robust optimization framework with a robust counterpart for convex optimization problems that lets the input vary within a specified range and selects the worst-case outcome (Ben-Tal and Nemirovski 1998). Also, there is the resampled frontier (Michaud 1998) which is constructed by resampling the input from a distribution and averaging over the resamplings’ optimization results.

Robust optimization approaches work well when all market information is quantified and incorporated in the optimization problem. But, despite efforts to incorporate information such as transaction costs, expert opinion and liquidity, an optimization problem remains a simplification of reality. Therefore, in practice, investors combine the optimal portfolio with additional information that was not or could not be incorporated. So, for investors, portfolio optimization is a tool that supports but does not replace their decision process. In this paper, we take this as the starting point for developing a robust optimization approach.

Near-optimal portfolios

Generally, the result of a portfolio optimization problem is an efficient frontier with optimal portfolios. Now, given an optimal portfolio on the efficient frontier, we construct a region of portfolios just below the efficient frontier as indicated by the shaded region in Figure 1. As shown in Chopra (1993) and Section 4, when optimal portfolios are sensitive to the optimization’s input parameters, these so-called near-optimal portfolios can have completely different weights than the optimal portfolio. The idea is that, for the investor, all near-optimal portfolios have satisfactory risk-return trade-offs and the investor can, as is already common practice, bring in additional arguments to select his preferred near-optimal portfolio.

To find the shaded region in Figure 1, we construct near-optimal portfolios \( w_0, \ldots, w_n \) that are far away from each other and show that any weighted average of these portfolios, i.e., any portfolio in their convex hull

\[
\text{Conv}(w_0, \ldots, w_n) = \left\{ \sum_{i=0}^{n} \theta_i w_i \left| \sum_{i=0}^{n} \theta_i = 1 \right. \right\},
\]

is near-optimal. We continue constructing near-optimal portfolios until their convex hull sufficiently covers the near-optimal portfolios represented by the shaded region in Figure 1. We
Figure 1 shows a mean-variance efficient frontier based on the statistics in Table 1, an optimal allocation (orange dot) and a shaded region with near-optimal portfolios. Table 1 shows statistics of monthly returns from January 1980 to December 1990 as reported by Chopra (1993) and portfolio weights of the orange dot in Figure 1.

will show that the region of near-optimal portfolios is more robust than the optimal portfolio \( w_0 \) itself. The intuitive understanding is that, with slightly different input parameters, the near-optimal region slightly changes in shape, but most near-optimal remain near-optimal. For example, the old optimum becomes near-optimal and one of the near-optimal portfolios becomes optimal. For the investor, his allocation becomes more robust, because no revision is needed when it remains near-optimal with slightly different input parameters.

3 Methodology

3.1 Constructing near-optimal portfolios

As discussed, the construction of near-optimal portfolios consists of the following steps:

1. start with an efficient frontier and an optimal portfolio \( w_0 \) on the frontier as in Figure 1;
2. specify the near-optimal region as indicated by the shaded region in Figure 1;
3. find the portfolio \( w_1 \) in the near-optimal region that is furthest away from \( w_0 \);
4. find the portfolio \( w_2 \) in the near-optimal region that is furthest away from \( \text{Conv}(w_0, w_1) \);
5. continue until \( \text{Conv}(w_0, \ldots, w_n) \) covers the near-optimal region up to a required precision \( \varepsilon \).

In this section, we specify these steps in more detail. Although near-optimal portfolios can be found below any efficient frontier, we assume for simplicity that the investor is interested in near-optimal mean variance portfolios. Thus, the efficient frontier is constructed by solving

\[
\begin{align*}
\min_w \lambda w^T \Sigma w - w^T \mu, \\
Aw &= b, \\
Gw &\leq h,
\end{align*}
\]

where the vector \( \mu \) contains the asset’s mean returns, \( \Sigma \) is their covariance matrix, \( A \) is a matrix representing together with the vector \( b \) the equality constraints, \( G \) is a matrix representing together with the vector \( h \) the inequality constraints and \( \lambda \geq 0 \) represents the investor’s risk.

<table>
<thead>
<tr>
<th>Stocks</th>
<th>Bonds</th>
<th>T-Bills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.323%</td>
<td>1.027%</td>
</tr>
<tr>
<td>Stdev.</td>
<td>4.793%</td>
<td>3.984%</td>
</tr>
</tbody>
</table>

Correlations:

| Stocks | 1.000 |
| Bonds  | 0.341 |
| T-Bills| -0.081|

Optimal alloc. | 58.1% | 22.8% | 19.1%
aversion. Also, equality constraints (2b) should at least enforce that all weights sum to one. Apart from this restriction, constraints (2b) and (2c) be chosen freely and, for example, be used to prevent short-selling or fix the allocation to certain asset classes. In Step 1, the investor solves mean variance optimization problem (2) for a number of risk aversion parameters $\lambda$ and obtains an efficient frontier denoted in Figure 1 by the solid blue line. From the frontier, the investor selects an optimal portfolio $w_0$ with an appropriate risk-return trade-off. In Step 2, the investor specifies a region $R$ of near-optimal portfolios that, compared to the optimal portfolio $w_0$, have an average return which is at most $\delta_\mu$ lower, i.e.,

$$w^T \mu \geq w_0^T \mu - \delta_\mu,$$

and have a variance which is at most $\delta_\Sigma$ higher, i.e.,

$$w^T \Sigma w \leq w_0^T \Sigma w_0 + \delta_\Sigma.$$

The near-optimal region $R(w_0, \delta_\mu, \delta_\Sigma)$ consists thus of all portfolios that satisfy (2b), (2c), (3) and (4), and, it is represented by the shaded region in Figure 1. It can be shown that the near-optimal region is convex which means that any weighted average of the near-optimal portfolios is near-optimal, see Appendix A.

Next, in Step 3, we find a portfolio $w_1$ in the region $R$ that is furthest away from $w_0$, i.e, $w_1$ satisfies:

$$w_1 = \arg \max_{w \in R(w_0,\delta_\mu,\delta_\Sigma)} \|w_0 - w\|,$$

where $\|\cdot\|$ indicates the Euclidean norm, i.e., the root of the sum of the components. Note when maximizing that the Euclidean norm favors large deviations in one component over small deviation in several components. More generally, as follows from Lemma A.1, once we found $i$ near-optimal portfolios $w_0, ..., w_{i-1}$, all portfolios in the convex hull $\text{Conv}(w_0, ..., w_{i-1})$, i.e., all weighted averages, are near-optimal. Therefore, in Step 4, we find $w_i$ by finding the portfolio furthest away from the convex hull of $w_0, ..., w_{i-1}:

$$w_i = \arg \max_{w \in R(w_0,\delta_\mu,\delta_\Sigma)} d(w, \text{Conv}(w_0, ..., w_{i-1})),

where the function $d$ indicates the distance to the convex hull:

$$d(w, \text{Conv}(w_0, ..., w_{i-1})) = \min_{w \in \text{Conv}(w_0, ..., w_{i-1})} \|w_i - w\|.$$

As indicated in Step 5, we continue constructing near-optimal portfolios until the constructed convex hull $\text{Conv}(w_0, ..., w_n)$ covers the near-optimal region $R(w_0, \delta_\mu, \delta_\Sigma)$ up to a required precision $\varepsilon > 0$, i.e., until

$$d(w_n, \text{Conv}(w_0, ..., w_{n-1})) < \varepsilon.$$

In particular, criterion (8) enforces there are no near-optimal portfolios outside of the convex hull of $w_0, ..., w_n$ that have allocation weights that differ more than $\varepsilon$ with the nearest portfolio in the convex hull. Therefore, in practice, we recommend choosing $\varepsilon$ equal to the absolute difference in allocation that is considered insignificant.

### 3.2 Support vector machines

Although Section 3.1 specifies how to construct the region of near-optimal portfolios, especially optimization problem (6) is difficult to solve since evaluation of its objective requires
solving optimization problem (7) to determine the distance of a portfolio to a convex hull. Optimization problem (6) can be simplified by applying theory on support vector machine (SVM) classification developed over time since Vapnik (1963).

In its easiest form, SVM-classification is a machine learning method that tries to separate two classes of points in space by a plane (Boser, Guyon, and Vapnik 1992). Figure 2 shows a two-dimensional example where the green points, representing portfolios $w_0$ to $w_{i-1}$, are separated in space from the orange point, representing the portfolio $w_i$, by the line $w^Tx + z = 0$; here $x$ is a vector and $z$ is a scalar. In SVM-classification, the separating line $w^Tx + z = 0$ maximizes its distance to the two classes which, as indicated in Figure 2, can be shown to equal $\frac{1}{\|x\|}$. As shown in (Boser, Guyon, and Vapnik 1992), the separating line of a SVM-classification problem can, when it exists, be found by solving a quadratic programming problem:

$$\min_{x, z} \|x\|,$$  

$$x^Tw_j + z \leq -1 \text{ for } j = 0, ..., i - 1,$$  

$$x^Tw_i + z \geq 1.$$  

As shown in Bennett and Bredensteiner (2000), Bennett and Campbell (2000), and Mavroforakis and Theodoridis (2006), the separation margin $\frac{1}{\|x\|}$ equals the distance between the convex hull $\text{Conv}(w_0, ..., w_{i-1})$ and the point $w_i$ as defined by (7). Optimization problem (6) can therefore be written as:

$$\min_{x, z, w_i} \|x\|,$$  

$$x^Tw_j + z \leq -1 \text{ for } j = 0, ..., i - 1,$$  

$$x^Tw_i + z \geq 1,$$  

$$Aw_i = b,$$  

$$Gw_i \leq h,$$  

$$w_i^T\mu \geq w_0^T\mu - \delta_\mu,$$  

$$w_i^T\Sigma w \leq w_0^T\Sigma w_0 + \delta_\Sigma,$$  

where constraints (10d) to (10g) enforce that $w_i$ is near-optimal, i.e., $w_i$ is in the region $R(w_0, \delta_\mu, \delta_\Sigma)$, and constraints (10b) and (10c) together with objective (10a) enforce that $w_i$ is at maximum distance of the convex hull $\text{Conv}(w_0, ..., w_{i-1})$.  

Figure 2: Graphical representation of a SVM-classification problem. The blue line separates the green portfolios from the orange portfolio and has maximal separation margin $\frac{1}{\|x\|}$. 

Conv($w_0, ..., w_{i-1}$)
Table 2: Near optimal portfolios and their centroid.

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Mean (%)</th>
<th>Std (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stocks</td>
<td>Bonds</td>
<td>T-Bills</td>
</tr>
<tr>
<td>W₀</td>
<td>58.1</td>
<td>22.8</td>
</tr>
<tr>
<td>W₁</td>
<td>7.7</td>
<td>85.0</td>
</tr>
<tr>
<td>W₂</td>
<td>50.3</td>
<td>0.0</td>
</tr>
<tr>
<td>W₃</td>
<td>73.7</td>
<td>0.0</td>
</tr>
<tr>
<td>W₄</td>
<td>46.1</td>
<td>52.7</td>
</tr>
<tr>
<td>W₅</td>
<td>62.6</td>
<td>27.6</td>
</tr>
<tr>
<td>W₆</td>
<td>27.8</td>
<td>71.4</td>
</tr>
<tr>
<td>W₇</td>
<td>55.0</td>
<td>40.6</td>
</tr>
<tr>
<td>C</td>
<td>46.2</td>
<td>39.6</td>
</tr>
</tbody>
</table>

Although optimization problem (10) is non-linear and not even convex due to inequality constraint (10c), it is solvable with standard optimization software such as the SciPy’s SLSQP and basin hopping solvers (Jones, Oliphant, Peterson, et al. 2001).

4 Example

4.1 Discussion

In this section, we continue with the Chopra (1993) example also shown in Figure 1. Under a no short selling constraint, we perform a mean variance optimization using the statistics in Table 1. This results in the efficient frontier (blue line) and the optimal portfolio $w₀$ (orange dot) indicated in Figure 3. The near-optimal region, indicated by the shaded region in Figure 1, consists of portfolios that, compared to the optimal portfolio (orange dot), have a return that at most 10% lower and a standard deviation is at most 10% higher. To find the near-optimal region, we apply the methodology in Section 3 and obtain portfolios $w₁$ to $w₇$. The convex hull of the portfolios $w₀,...,w₇$ covers the near-optimal region up to a precision $ε = 0.01$ implying that no near-optimal portfolios that with allocation differences larger than 1% exist outside of the convex hull, see equation (8).

Chopra (1993) studies (almost) the same near-optimal region with the purpose to show that near-optimal portfolios can have completely different weights. There are two important differences to note. First, for the purpose in Chopra (1993), a different definition of the near-optimal region is used: portfolios that, compared to the optimal portfolio, provide 90% of the average return for less than 90% of the standard deviation are left out, i.e., the lower left corner of the near-optimal region indicated in Figure 3 is left out. For our purpose, however, we do consider these portfolios near-optimal because there are portfolios considered near-optimal with the same average return and a higher standard deviation. Regardless of this difference in definition, we verified that, as is implied by the difference, all near-optimal portfolios constructed in Chopra (1993) are contained in the convex hull of $w₀,...,w₇$.

Second, Chopra (1993) constructs near-optimal portfolios through a grid search, i.e., try all possible portfolios, and searches for near-optimal portfolios with the highest upward and
downward deviation in one asset class. Because there are three assets, this results in six near-optimal portfolios in addition to the optimal portfolio $w^*_0$ that, as intended by Chopra (1993), differ completely in weights. It, however, can be verified that the convex hull of the near-optimal portfolios found in Chopra (1993) does not completely cover the near-optimal region. For example, portfolios $w^*_4$ and $w^*_5$ are near-optimal in the definition of Chopra (1993), but cannot be written, also not approximately, as a weighted average of the near-optimal portfolios reported in Chopra (1993). This shows that constructing portfolios with the highest upward and downward deviation is not suitable to find the complete near-optimal region. Additionally, a grid search algorithm is, contrary to the methods presented here, only feasible in very low dimensions.

Since optimization problem (10) is not convex, we performed two consistency checks to ensure the convex hull of the near-optimal portfolios $w^*_0, \ldots, w^*_7$ indeed covers the near-optimal region. First, as noted, we verified that all near-optimal portfolios constructed in Chopra (1993) are contained in the convex hull of by $w^*_0, \ldots, w^*_7$. Second, we also verified that all portfolios both on the efficient frontier and in the near-optimal region are contained in the convex hull of $w^*_0, \ldots, w^*_7$. Together, this gives sufficient confidence in the convergence and accuracy of numerical solvers.

Note that Figure 3 might wrongly give the impression that the near-optimal portfolios of which the convex hull covers the near-optimal region should lie on the boundary of the region indicated by the shaded region in Figure 3. It can easily be shown that this is not the case. For example, when the near-optimal region is increased to portfolios with a return of at least 1% and a variance of at most 4%, the first near-optimal portfolio found has a 100% allocation to treasury bills and does not lie on the boundary of the region indicated by the shaded region in Figure 3.

4.2 Selecting a preferred near-optimal portfolio

Once the near-optimal region is covered by the convex hull of the near-optimal portfolios $w^*_0, \ldots, w^*_7$, the investor is free to select a preferred portfolio from this region. Without further information, the centroid $c$ of the near optimal portfolios $w^*_0 \ldots w^*_7$, indicated by the green dot in Figure 1, can be a good default choice:

$$c = \frac{1}{8} \sum_{i=0}^{7} w^*_i. \quad (11)$$

We will show that the centroid is robust in two ways. First, the centroid is less sensitive to changing input parameters than portfolios on the frontier. Second, because the centroid is not on the boundary of the convex hull, it is expected to remain near-optimal with slightly different input parameters.

Although the centroid portfolio can be a good default choice, preferably, the investor brings in additional arguments to select a preferred near-optimal portfolio. In light of the sensitivity problem, these arguments should be additional to the risk-return statistics in Table 1, e.g., selecting the portfolio with the highest sharp ratio would not suffice. As an example, suppose the investor prefers, for whatever reason, not to invest in bonds. It follows from Table 2 that, in that case, any weighted average of portfolio $w^*_2$ and $w^*_3$ would suffice: such a portfolio is near-optimal and has no allocation to bonds. When the investor currently owns 60% equity and 40% cash which is to be invested in either stocks or treasury bills, he could, for example, choose to leave his current exposure to equity intact and to buy treasury bills with his cash.
Figure 4 and 5 show, as a function of the sample size $N$, average percentage of overlap between the original and perturbed convex hull (yellow), the average turnover between the original and the perturbed mean variance optimal portfolio (orange) and the average turnover between portfolios in the original convex hull and the closest portfolio in the perturbed convex hull (blue). In addition, Figure 5 shows the original and the perturbed centroid (green). And also, the average turnovers between the closest portfolio in the perturbed convex hull and: the original convex hull (blue), the original mean variance optimal portfolio (orange dotted), and the original centroid portfolio (green dotted).

4.3 Robustness of the preferred near-optimal portfolio

To investigate the robustness of the near-optimal portfolio method, we view the mean and covariance matrix presented in Table 1, i.e., the optimization’s input parameters, as estimated on a sample of size $N$. To perturb the optimization’s input, we draw a sample of size $N$ from a normal distribution with mean and covariance as in Table 1 and estimate a perturbed mean and covariance matrix. So, the larger the sample size the closer the perturbed means and covariance matrices are to original ones in Table 1. For each perturbed mean and covariance matrix, we determined the perturbed optimal mean variance portfolio, near-optimal portfolios, convex hull and centroid portfolio.

Figures 4 and 5 shows the average robustness using 100 samples for each sample size. In Figure 4, the yellow dots and their trend line show for each sample size the average percentage overlap using 100 sample of size $N$ between the original and perturbed convex hulls. The increase of the yellow trend shows that, with perturbed input parameters, roughly 60% to 90% of the near-optimal portfolios remains near-optimal and, consequently, no new investment advice is required.

The comparison with mean variance optimization can best be made with the turnover measure:

$$T(u, v) = \frac{1}{2} \sum_i |u_i - v_i|,$$

where the sum is taken over all the components of the vectors. The turnover measure can be interpreted as the fraction of the portfolio $u$ that has to be sold and reinvested to obtain portfolio $v$. In Figures 4 and 5, the orange dots and their trend line show that the average turnover between the mean variance optimal and perturbed mean variance optimal portfolio ranges from roughly 10% to 25%. The average turnover is significantly decreased with the near-optimal portfolio method: the blue dots and their trend line in Figures 4 and 5 indicate that on average less than 5% of the portfolio has to be sold and reinvested to obtain a near-optimal portfolio when input parameters are re-estimated.

$^1$Equivalently, we could have drawn perturbed means and covariance matrices from a normal-inverse-Wishart distribution which is the conjugate prior to the multivariate normal distribution.
Figure 5 shows for two typical near-optimal portfolios, the mean variance optimal portfolio and centroid portfolio, that their robustness significantly increases when their robustness is measured w.r.t. the near-optimal region. At least for these two near-optimal portfolios, the near-optimal method achieves its increase in robustness, because it measures robustness with respect to a region of near-optimal portfolios instead of with respect to a single optimal portfolio. In other words, the increase in robustness is a free advantage when all near-optimal portfolios are considered valid investment decisions.

5 Conclusion

We have shown how to construct a region of near-optimal portfolios below the efficient frontier. Also, we discussed how this methodology is both robust and is designed to support rather than replace the investor’s investment decision process.

There are several directions that can be explored in future research. First, the methodology can be applied to more computationally intensive optimization objectives such as mean-CVaR optimization. Second, the robustness w.r.t. estimates for the mean and covariance matrices can be explored separately. Also, the robustness can be compared with other robust optimization approaches such as resampling. And finally, the method’s integration with the investment decision process can be further explored.

6 Acknowledgments

We would like to thank Dr. O.W. van Gaans of Leiden University for contributing to this research and supervising a master thesis on this topic. Additionally, we would like to thank our colleges Marnix Engels and Loranne van Lieshout for their valuable feedback.

A Convexity of the near-optimal region

Lemma A.1. The near-optimal region $R(w_0, \delta_\mu, \delta_\Sigma)$ consisting of all portfolios that satisfy (2b), (2c), (3) and (4) is convex.

Proof. When given two near-optimal portfolios $u, v \in R(w_0, \delta_\mu, \delta_\Sigma)$, we have to show that any weighted average $w = tu + (1-t)v$ is also near-optimal, i.e., $w$ satisfies (2b), (2c), (3) and (4) for all $0 \leq t \leq 1$. First, since $u$ and $v$ satisfy (2b), it directly follows that $w$ satisfies (2b). Also, inequality constraints (2c) and (3) follow directly. That $w$ satisfies inequality constraints (4) follows from convexity of the left hand side of (4) and applying Jensen’s inequality:

$$w^T \Sigma w \leq tu^T \Sigma u + (1-t)v^T \Sigma v \leq t(w_0^T \Sigma w_0 + \delta_\Sigma) + (1-t)(w_0^T \Sigma w_0 + \delta_\Sigma) = w_0^T \Sigma w_0 + \delta_\Sigma.$$

References


DOI: 10.1287/moor.23.4.769. (Cit. on p. 1).

DOI: 10.1287/opre.1120.1115. (Cit. on p. 1).

DOI: 10.1145/130385.130401. (Cit. on p. 4).

DOI: 10.1073/pnas.0904287106. (Cit. on p. 1).


DOI: 10.3905/joi.2.3.51. (Cit. on pp. 1, 2, 5, 6).

DOI: 10.1080/14697680701422089. (Cit. on p. 1).

DOI: 10.1287/mnsc.1080.0986. (Cit. on p. 1).


DOI: 10.1016/j.orrp.2015.06.003. (Cit. on p. 1).


DOI: 10.1109/TNN.2006.873281. (Cit. on p. 4).


