A Sharper Angle on Optimization

Maxim Golts and Gregory Jones[*]

The classical mean-variance optimization takes expected returns and variances and produces portfolio positions. In this paper we discuss the *direction* and the *magnitude* of the positions vector separately, and focus on the former. We quantify the distortions of the mean-variance optimization process by looking at the *angle* between the vector of expected returns and the vector of optimized portfolio positions. We relate this angle to the condition numbers of the covariance matrix and show how to control it by employing robust optimization techniques. The resulting portfolios are more intuitive and investment-relevant, in particular with lower leverage of the "noise" alphas at the expense of lower ex-ante Sharpe Ratio.

First version: May 8, 2009
This version: November 15, 2009

[*] Golts: maxgolts@gmail.com, Jones: Gregory.Jones@gmo.com. We are grateful to Brian Chen, Neil Constable, David Cowan, Chris Darnell, Dan diBartolomeo, Andy Gossard, Ben Inker, Jakub Jurek, Mark Mueller, Dessi Pachamanova, Nehal Patel, Gordon Ritter, Melanie Rudoy, and Sam Wilderman for useful discussions, and to JPMorgan for the Global Bond Index data.

1 Introduction

In spite of the simplicity and elegance of the long-established mean-variance optimization introduced by Markowitz (1952, 1959, 1987), many portfolio and risk managers are still wary of it. One of the frequent complaints about the optimization heard from fellow practitioners is that the portfolio positions it recommends are non-intuitive, bearing little resemblance to the manager's expected returns or "alphas." In this article we introduce the concept of *alpha-weight angle* in an attempt to quantify how different the alphas and optimized positions really are. If the angle is not "too big," the positions reasonably reflect the alphas, but if the angle is close to 90 degrees, the alphas and positions are nearly orthogonal and the portfolio has little to do with the manager's investment insights.

We show that if the covariance matrix is very ill-conditioned or "close to degenerate," the alpha-weight angle could be very large, whereas well-conditioned matrices ensure the angle remains within tight bounds. In the midst of a crisis, spiking volatilities and higher correlations will make a covariance matrix more ill-conditioned. Our analysis provides another way to see that mean-variance optimization can result in "error maximization," or excessive *leverage* of the "noise" alphas and the overall portfolio. This problem could be especially pronounced with a high ex-ante Sharpe Ratio, forecasted with historical covariance estimates during times of great instability and uncertainty. In practice the ex-ante Sharpe Ratio could be overestimated, and excessively high leverage could be very dangerous.

How can this problem be mitigated? The literature is very extensive, see for example Jorion (1986, 1992), Black and Litterman (1992), Chopra and Ziemba (1993), Michaud (1989, 1998), Jacobs, Levy, and Markowitz (2006) and references therein.

Our idea here is to separate the *direction* and *magnitude* of the positions vector.[1] We suggest making the covariance matrix better-conditioned or "more regular" if needed, for the "direction" part of the optimization, thereby constraining the alpha-weight angle. One way to do this is to introduce certain robust optimization procedures, which in particular reduce the ex-ante Sharpe Ratio but also reduce the leverage of the "noise" alphas. The magnitude of positions should be determined separately, derived from restrictions on the overall leverage, tracking error, return target, or other possibly forward-looking considerations unrelated to historical covariance estimates, such as exogenous market events, government policies, or liquidity constraints.

Overall, we suggest here a 3-step leverage control process: first, control the leverage of the optimization *inputs*[2]; second, constrain the excessive leverage of small alphas by aligning the *directions* of the alphas and the weights vectors; and third, set the total portfolio leverage by scaling the *magnitude* of the weights vector.

2 Classical Mean-Variance Optimization: Direction and Magnitude

Consider N assets with expected single-period returns (alphas) $\alpha_1,\ldots,\alpha_N$ and the $N \times N$ covariance matrix[3] denoted by Σ. The goal of any mean-variance procedure is to transform alphas and covariances into portfolio positions (weights) $w_1,\ldots,w_N$ while optimizing the portfolio return/risk tradeoff.

2.1 Four optimization problems

Let us treat the collections of alphas and weights for N assets as N-dimensional vectors

$$\alpha = (\alpha_1, \ldots, \alpha_N) \qquad \text{and} \qquad w = (w_1, \ldots, w_N). \tag{1}$$

It will be convenient to use the vector notation throughout this paper. First, denote by $|\alpha| = \sqrt{\alpha'\alpha}$ and $|w| = \sqrt{w'w}$ the lengths of the vector α and w respectively, and denote by $\hat{\alpha} = \alpha/|\alpha|$ and $\hat{w} = w/|w|$ the corresponding unit vectors.

The expected return and expected variance of the portfolio are respectively:

$$r_P = \sum_{i=1}^{N} \alpha_i w_i = \alpha' w \qquad \text{and} \qquad \sigma_P^2 = \sum_{i,j=1}^{N} w_i \Sigma_{ij} w_i = w'\Sigma w, \tag{2}$$

where Σ_{ij} are the elements of the matrix Σ. In **Table 1**, we list four classical Markowitz-type optimization problems.

Table 1. Classical Optimization Problems

Number and Name	**Problem**	**Solution**
I. "Risk Constraint"	$\max_w r_P$ <u>subject to</u> $\sigma_P^2 \le \sigma_0^2$	$w = \dfrac{\sigma_0}{\sqrt{\alpha'\Sigma^{-1}\alpha}} \Sigma^{-1}\alpha$
II. "Return Constraint"	$\min_w \sigma_P^2$ <u>subject to</u> $r_P \ge r_0$	$w = \dfrac{r_0}{\alpha'\Sigma^{-1}\alpha} \Sigma^{-1}\alpha$
III. "Risk Aversion"	$\max_w r_P - \lambda \sigma_P^2$	$w = \dfrac{1}{2\lambda} \Sigma^{-1}\alpha$
IV. "Sharpe Ratio"	$\max_w \dfrac{r_P}{\sqrt{\sigma_P^2}}$	$w = \dfrac{w'\Sigma w}{\alpha' w} \Sigma^{-1}\alpha$

Notes: We consider portfolios employing leverage and long/short positions, and hence do not have constraints of the type $\sum_{i=1}^{N} w_i = 1$ or $w_i \ge 0$ for all $1 \le i \le N$. In problem III, λ is the "risk aversion" parameter. For simplicity, we assume that the risk-free rate is zero in problem IV.

2.2 Discussion: separate direction and magnitude

In all four problems listed in **Table 1**, the directions of the solutions are the same, $w \propto \Sigma^{-1}\alpha$. The magnitude is obtained from the tracking error constraint for Problem I, the return target for Problem II, and the risk aversion parameter λ for Problem III. Problem IV is magnitude-independent.

We propose to separate the mean-variance optimization process into two steps:

1. Investment "direction" defined by the unit vector $\hat{w}$.
2. Investment "magnitude" defined by a norm of w, which could be leverage $\sum_i |w_i|$, or tracking error $\sqrt{w'\Sigma w}$. Once the direction is chosen, all magnitude decisions are simultaneously scalable, so if $\hat{w}$ is fixed and the tracking error is increased by a factor of 2, the leverage will also increase by a factor of 2.

We can go further and consider the specification of the direction and magnitude to be separate, independent problems. In particular we can use two different covariance matrices or even two completely different procedures to establish the direction and magnitude of the portfolio positions.

3 Angle between Alphas and Weights

3.1 Definitions

The mean-variance optimized portfolio weights could be quite different from the original alphas. We propose to quantify this difference by looking at the *angle* ω between the vectors of alphas and weights:

$$\cos(\omega) = \frac{\alpha' w}{\sqrt{\alpha'\alpha}\sqrt{w'w}}. \tag{3}$$

If this angle is too big, the weights probably bear little resemblance to the original alphas. An estimate on this angle can be obtained from the spectral properties of the covariance matrix Σ, which we assume to be positive definite, so it can be presented as

$$\Sigma = Q'\mathrm{diag}(\theta_1^2, \theta_2^2, \ldots, \theta_N^2)Q, \tag{4}$$

where Q is an orthogonal matrix and $\theta_1^2 \geq \theta_2^2 \geq \ldots \geq \theta_N^2$ are the eigenvalues of Σ in decreasing order. The rows of Q denoted by $q_1, q_2, \ldots, q_N$ are the eigenvectors (principal components) of Σ. Denote $\theta_{\max}^2 = \theta_1^2$ and $\theta_{\min}^2 = \theta_N^2$, and $q_{\max} = q_1$ and $q_{\min} = q_N$.

3.2 Worst-case angles

As we have seen above, the direction of the optimized weights is given by $w \propto \Sigma^{-1}\alpha$. Then the quantity $\cos(\omega)$ is bounded from below as follows:

$$\cos(\omega) = \frac{\alpha'\Sigma^{-1}\alpha}{\sqrt{\alpha'\alpha}\sqrt{\alpha'\Sigma^{-2}\alpha}} \geq \frac{\theta_{\max}\theta_{\min}}{(\theta_{\max}^2 + \theta_{\min}^2)/2} = 2\sqrt{\frac{\kappa}{(\kappa+1)^2}}, \tag{5}$$

where $\kappa = \theta_{\max}^2/\theta_{\min}^2$ is the *condition* number of the matrix Σ, and the equality is achieved if and only if

$$\alpha = \pm\theta_{\max}q_{\max} \pm \theta_{\min}q_{\min} \qquad \text{and} \qquad w \propto \pm q_{\max}/\theta_{\max} \pm q_{\min}/\theta_{\min}. \tag{6}$$

We will denote

$$\mathrm{mmd}(\Sigma) = \frac{\theta_{\max}\theta_{\min}}{(\theta_{\max}^2 + \theta_{\min}^2)/2} \tag{7}$$

and call this quantity the *minimax degeneracy* number of the matrix Σ. Further discussion of matrix degeneracy measures is given in the **Appendix**.

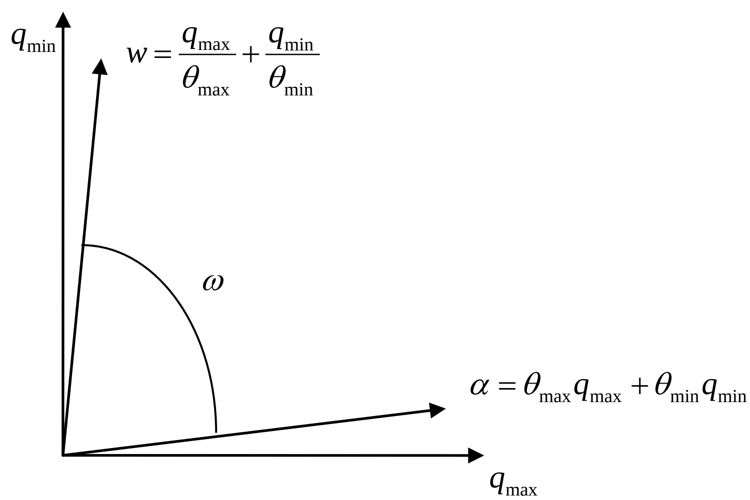
3.3 The worst case angle discussion: nearly orthogonal investing

The worst case is illustrated by **Figure 1**. This situation is problematic because the portfolio is largely invested along the lowest volatility principal component $q_{\min}$ and the alpha component in this direction is also very small and could be subject to an estimation error or noise in the data. However, this small alpha $\theta_{\min}$ is divided by an even smaller number $\theta^2_{\min}$ and is thus levered up by the optimization. In general, looking at

$$\text{Expected Portfolio Return} = r_P = \alpha' w = |w||\alpha|\cos(\omega), \qquad (8)$$

we see that when $\cos(\omega) \approx 0$, the choice of investment direction $\hat{w} = w/|w|$ does not yield a significant expected return. Indeed, a small error in the estimation of α may yield a *negative* $\cos(\omega)$ and hence *negative* expected return. The size of the expected portfolio return arises only from an overextended leverage in the $\hat{w}$ direction. This is not a sound investment practice in any setting involving uncertainty. From (8) we see that the problem of overleverage is directly tied to the problem of small $\cos(\omega)$. Another way to look at this: in the presence of small covariance eigenvalues, standard mean-variance optimal weights can be extremely sensitive to changes in the input alphas. In this sense the mean-variance optimization can inappropriately amplify a low-confidence investment, or to use the colloquialism, "maximize the error." This phenomenon is extensively studied and discussed in the literature, see for example Michaud (1989), Jorion (1992), and Kritzman (2006).

Figure 1. The worst case angle



The robust optimization discussed below takes the estimation errors into account and helps to mitigate the "worst case" problem.

3.4 A remedy: "shrink the covariance matrix to the identity"

How can we make the alphas and optimized weights more aligned, that is, make sure that the angle between α and w is not too big? We could modify the covariance matrix, increasing its minimax degeneracy number, and use the modified matrix for optimization. A simple and popular way to do that is to consider the following parametric family of "shrunk to identity" matrices

$$\Sigma(t) = t\frac{\mathrm{tr}\Sigma}{N} I_N + (1-t)\Sigma \qquad (9)$$

for t between 0 and 1. This transformation averages covariance eigenvalues with their arithmetic mean, and hence avoids the division-by-small-numbers problem and leads to more acute angles. In particular, since $\mathrm{mmd}(I_N) = 1$ and $\mathrm{mmd}(\Sigma(t))$ is a continuous function of t, there exists $\tilde{t}$ between 0 and 1 such that $\mathrm{mmd}(\Sigma(\tilde{t}))$, and thus $\cos(\omega)$, is greater than any desired prespecified level.

4 Robust Optimization and the Alpha-Weight Angle

4.1 Four robust optimization problems

In this section we will show how to constrain the alpha-weight angle using a robust optimization procedure. In the robust worldview, the alphas are presumed to be uncertain,[4] with the vector α lying in some uncertainty region U_α, see for example Goldfarb and Iyengar (2003), Tütüncü and Koenig (2004), Ceria and Stubbs (2006), Fabozzi, Kolm, Pachamanova, and Focardi (2007). We chose a portfolio to maximize utility in a worst-case scenario for realization of $\alpha \in U_\alpha$. The robust versions of the four classical optimization problems are listed in **Table 2**.

Table 2. Robust Optimization Problems

Number	Problem
I	$\max\limits_{w} \min\limits_{U_\alpha} r_P$ <u>subject to</u> $\sigma_P^2 \le \sigma_0^2$
II	$\min\limits_{w} \sigma_P^2$ <u>subject to</u> $\min\limits_{U_\alpha} r_P \ge r_0$
III	$\max\limits_{w} (\min\limits_{U_\alpha} r_P - \lambda \sigma_P^2)$
IV	$\max\limits_{w} \min\limits_{U_\alpha} \frac{r_P}{\sqrt{\sigma_P^2}}$

4.2 Spherical uncertainty region for alphas[5]

Suppose now that the uncertainty region U_α is a ball centered at α with the radius $\chi|\alpha|$ for χ between 0 and 1. This corresponds to a one-sigma neighborhood under a Bayesian prior of an uncertain α distributed normally about the estimated α, with $\sigma = \chi|\alpha|$, see Meucci (2005) for a review of Bayesian methods[6]. Then

$$\min_{U_\alpha} r_P = \alpha' w - \chi |\alpha| |w| = |\alpha| |w| (\cos(\omega) - \chi). \tag{10}$$

(If $\chi > 1$, then $\min\limits_{U_\alpha} r_P \le 0$.) The optimization function then involves both the usual expected returns term, and a regularization term (whose relative strength is controlled by χ), which penalizes angles for being less acute.

4.3 Robust weights: Direction

The robust problems in **Table 2** have no closed form solutions in general,[7] but we can express $\hat{w} = w / |w|$ as follows:

$$\hat{w} = (\chi I_N + \frac{\cos(\omega) - \chi}{\hat{w}'\Sigma\hat{w}}\Sigma)^{-1}\hat{\alpha}, \tag{11}$$

where $\hat{\alpha} = \alpha / |\alpha|$. It is easy to see that for $\hat{w}$ as above

$$\hat{\alpha}'\hat{w} = \cos(\omega) \geq \chi, \tag{12}$$

regardless of Σ or α. So all four robust optimization problems allow us to constrain the angle ω to be between 0 and $\cos^{-1}(\chi)$. In the **Appendix** we improve on (12) in a particular case. Note that the robust covariance matrix family

$$\Sigma(\chi, \alpha) = \chi I_N + \frac{\hat{\alpha}'\hat{w} - \chi}{\hat{w}'\Sigma\hat{w}}\Sigma \tag{13}$$

for χ between 0 and 1 regularizes the matrix Σ in a way similar[8] to (9). The effective parameter t coming from the robust problem is implicitly dependent on Σ and α, and so the matrix $\Sigma(\chi, \alpha)$ is "shrunk to the identity matrix." In the simple example below, we compare our "sharper angle optimization" shrinkage to the shrinkage of the prescription (5) and (9), which will modify Σ to protect against a hypothetical worst case α. Notice that the less we trust our alphas (the bigger the χ is), the more we force our optimized weights to be closer to them. This may seem counterintuitive; the explanation is that we don't want our optimization model to be a "garbage in, *levered* garbage out" black box.

4.4 Robust Weights: Magnitude

Once the unit vector $\hat{w}$ is found, the magnitude of w can be determined by a completely separate procedure, from desired leverage, tracking error, or target return constraints. The magnitude for Robust Problems I, II, and III is given by:

$$|w|_{\mathrm{I}} = \frac{\sigma_0}{\sqrt{\hat{w}'\Xi\hat{w}}}, \qquad |w|_{\mathrm{II}} = \frac{r_0}{|\alpha|(\cos(\omega) - \chi)}, \qquad |w|_{\mathrm{III}} = \frac{1}{2\lambda} \tag{14}$$

respectively. We have used a different notation Ξ for the covariance matrix here to emphasize that it could be different from Σ. Robust Problem IV only specifies $\hat{w}$, it does not have any magnitude restrictions.

5 Example: A Tale of Three Managers

5.1 Managers A, B, and C

As an illustration of the alpha-weight angle concept introduced in this paper, we will investigate three fictitious absolute return managers operating in the global bond space in 2006-08.

All three managers are trading monthly and are US-based; all target 5% tracking error; all run long/short portfolios of seven assets proxied by JPMorgan Global Bond Indices for the G7 countries (Canada, Germany, France, UK, Italy, Japan, US); and all have the same two-factor alpha model

$$\alpha = \beta_1 F_1 + \beta_2 F_2 + \varepsilon, \tag{15}$$

where $\beta_1 = \beta_2 = 1$, F_1 is the yield of JPMorgan Global Bond Indices for each country, and F_2 is the "flight to quality" factor adjusting alphas for the largest and most liquid markets in Japan and US up by 100 basis points, German alphas up by 50 basis points, and the rest down by 50 basis points.

Manager A's positions are directly proportional to the alphas above, scaled by 2.5 to roughly target the 5% tracking error.

Manager B is using an unconstrained Markowitz optimization with the covariance matrix estimated from the rolling monthly returns of the JPMorgan Global Bond Indices over the previous year. The covariance matrix used by manager B on 11/3/2008 is given in **Table 3**.

Manager C is starting with the same historically estimated covariance matrix as Manager B, but is then using our robust optimization procedure with $\chi = \cos(\pi/3) = 0.5$ thus ensuring that the angle between alphas and weights cannot be more than $\pi/3 = 60^\circ$. We implement the optimization routine using MATLAB's **fmincon** function. The 5% tracking error target determines the magnitude of the weights.

Table 3. Volatilities and Correlations of Global Bonds on 11/3/2008

	Canada	Germany	France	UK	Italy	Japan	US
Volatilities	15.81%	12.62%	12.61%	12.96%	12.83%	15.33%	5.69%
				Correlations			
Canada	100%						
Germany	63%	100%					
France	63%	**99.92%**	100%				
UK	70%	86%	86%	100%			
Italy	65%	**98.89%**	**99.32%**	87%	100%		
Japan	-24%	31%	29%	8%	23%	100%	
US	6%	37%	36%	18%	33%	73%	100%

Note: the volatilities and correlations are estimated from rolling 21-business-day returns of JPMorgan Global Bond Indices over the trailing year.

5.2 Performance

First let's consider the performance of the three managers from 12/31/2005 to 6/29/2007, or during the last 18 months of the low risk premia period, see **Figure 2.**[9] Manager B is leading the race at this point and is perhaps under the impression that the optimization is providing him/her with an edge over the competitors.

Now, let's move to the global crisis period, which we would say started roughly on 7/1/2007. Over the next 18 months to the end of 2008, and particularly in Q4 2008, Manager B has dramatically underperformed Managers A and C, see **Figure 3**.

Figure 2. Performance of Managers A, B, and C during the last 18 months of the low risk premia period

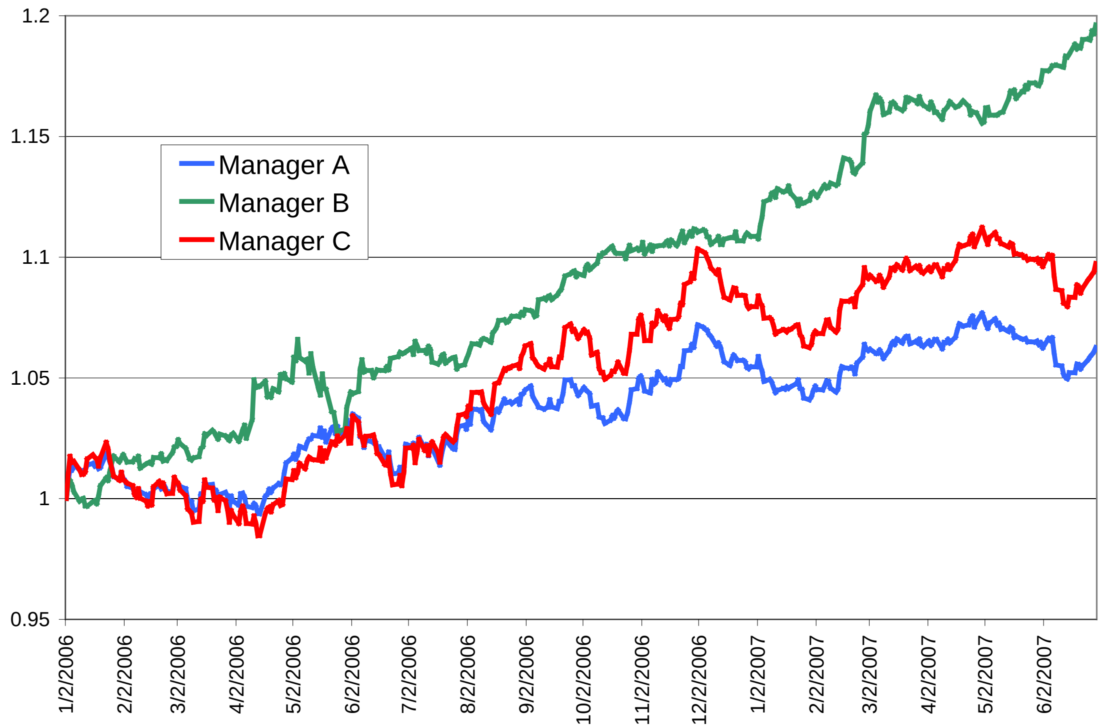
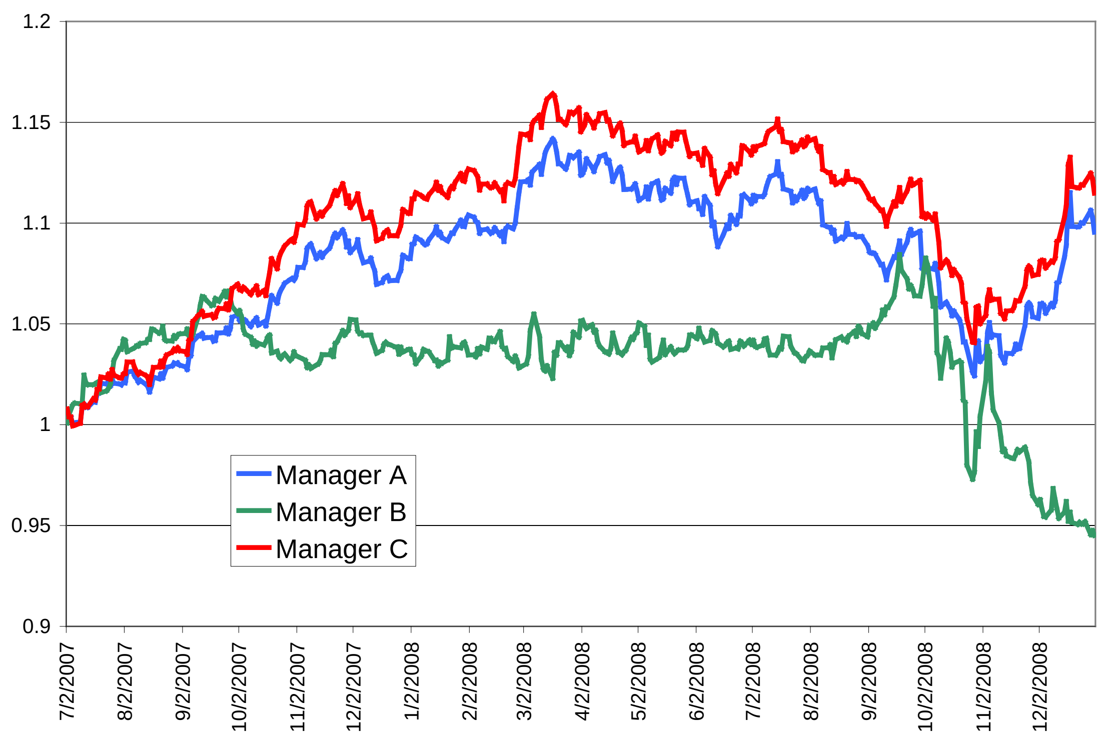


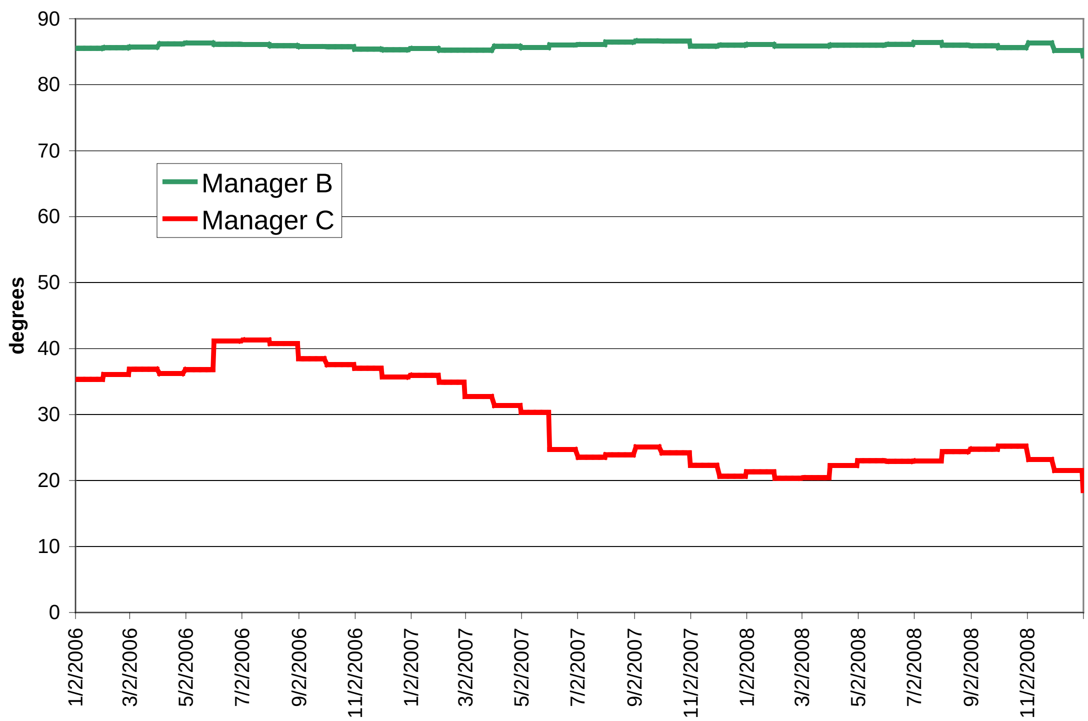
Figure 3. Performance of Managers A, B, and C during the crisis period



5.3 Angles

What happened? Let's look at the angles between the positions of Managers A, B, and C and their alphas (investment insights). Manager A is trading the rescaled alphas, so the angle in question is zero. **Figure 4** depicts the angles for Managers B and C.

Figure 4. Alpha-weight angles for Managers B and C in 2006-08



The angle between Manager's B weights and his/her alphas is about 85°. So, *the positions have little to do with the investment goals, and the performance is purely accidental*. In fact, Manager B, being long German bonds and short French bonds, is investing into the seventh, least volatile principal component of the covariance matrix. In November 2008, he/she increases the leverage for this spread trade precisely as the French bonds (which had underperformed the German bonds up to that date) begin to outperform German bonds. This is a typical example of an unrestricted optimizer going into spreads of highly correlated assets with a small difference in alphas.[10]

By contrast, Manager C's angle is optimized to be less than 60°, so the positions are much better aligned with alphas, see **Table 4**. (In fact, our optimization procedure will always deliver an angle strictly sharper than "requested" via χ: the solution is always an interior one. The more degenerate the covaraince matrix is, the less it is "trusted" by our robust optimization and the closer the optimized weights are to the alphas, see more examples below.)

Table 4. Positions of Managers A, B, and C as of 11/3/2008

	GBI Yield	Flight to Quality Factor	Manager A	Manager B	Manager C
Canada	3.89%	-0.50%	8.47%	-4.88%	5.52%
Germany	3.85%	0.50%	10.86%	761.31%	8.24%
France	4.13%	-0.50%	9.08%	-985.77%	4.75%
UK	4.50%	-0.50%	10.00%	6.51%	7.97%
Italy	4.93%	-0.50%	11.07%	218.07%	8.69%
Japan	1.46%	1.00%	6.16%	-6.33%	4.75%
US	3.62%	1.00%	11.54%	32.07%	19.35%

Notes: GBI Yield is the yield of JPMorgan Global Bond Indices

6 More Examples: Sharpe Ratio vs Leverage

We will consider in full detail the following toy examples with just two assets.

6.1 "Big" alpha-weight angle

Assume that the two assets we deal with have expected returns and estimated volatilities as in **Table 5.**

Table 5.

	Expected Return	Volatility	Ex-ante Sharpe Ratio
Asset 1	10%	10%	1.00
Asset 2	1%	1%	1.00

Assume further that a hedge fund manager is trying to achieve a 20% tracking error. Compare and contrast the optimized weights, leverage and alpha-weight angles for

1. Classical mean-variance optimization.
2. Optimization with the covariance matrix "shrunk" according to the prescription (9) , so that the minimax degeneracy number of the shrunk matrix is 0.5, guaranteeing the alpha-weight angle to be less than 60° via (5).
3. Our "Sharper Angle Optimization" with $\chi = \cos(\pi/3) = 0.5$.

The results are presented in **Table 6.**

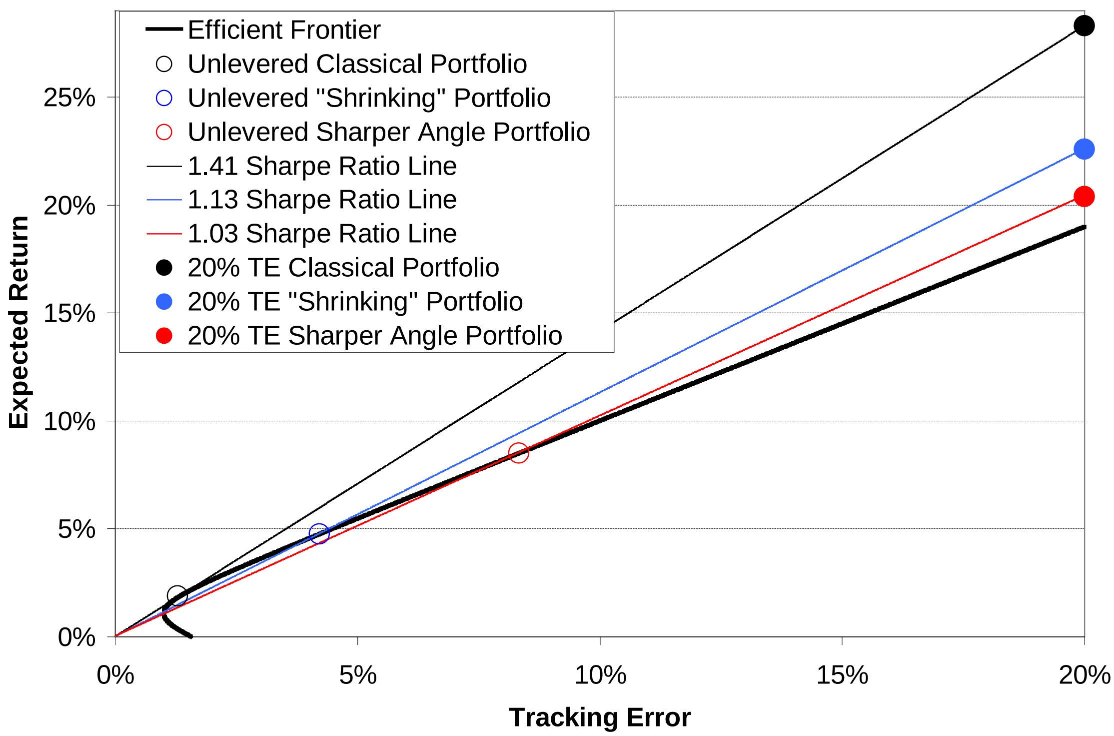
Table 6.

	Classical	Shrinking To Identity	Sharper Angle
Weight of Asset 1	141%	198%	200%
Weight of Asset 2	1414%	276%	40%
Leverage	1556%	474%	240%
Tracking Error	20%	20%	20%
Ex-ante Sharpe	1.41	1.13	1.02
Alpha-Weight Angle	79	49	6

Notes: the Alpha-Weight Angle is expressed in degrees.

In the classical case Asset 2 is levered over 14 times and the overall leverage is over 15.5 times. The "shrinking" optimization reduces both the leverage of Asset 2 and the overall portfolio. Our Sharper Angle optimization "distrusts" the close-to-degenerate covariance matrix and reduces the leverage numbers even further, aligning the alphas and weights much better than the "requested" 60°. We pay for the leverage reduction with lower ex-ante Sharpe Ratio. We believe that in practice this is a good tradeoff, since ex-ante Sharpe Ratio is a *forecast* and is subject to (over)estimation error, whereas leverage is a *fact*, and a clear and present danger. The Sharpe Ratio vs leverage tradeoff for the three optimized portfolios is illustrated in **Figure 5**: the *higher the ex-ante Sharpe Ratio, the more leverage is needed in order to achieve the desired tracking error target*.

Figure 5. Sharpe Ratio and Leverage for the Three Optimized Portfolios



Notes: The Efficient Frontier of the two assets presented in **Table 5**. "TE" stands for "Tracking Error" and "Shrinking" stands for Shrinking to Identity according to (9).

6.1 "Reasonable" alpha-weight angle

Assume that the two assets we deal with have expected returns and estimated volatilities as in **Table 7.**

Table 7.

	Expected Return	Volatility	Ex-ante Sharpe Ratio
Asset 1	1%	10%	0.10
Asset 2	1%	1%	1.00

The three optimization procedures yield the results reported in **Table 8.**

Table 8.

	Classical	Shrinking To Identity	Sharper Angle
Weight of Asset 1	20%	117%	63%
Weight of Asset 2	1990%	1625%	1897%
Leverage	2010%	1741%	1961%
Tracking Error	20%	20%	20%
Ex-ante Sharpe	1.00	0.87	0.98
Alpha-Weight Angle	44	41	43

Notes: the Alpha-Weight Angle is expressed in degrees.

In this case the angle between alphas and Markowitz-optimized weights is a "reasonable" 45°, already sharper than the "requested" constraint of 60°. Thus, neither Shrinking to Identity, nor Sharper Angle procedures change the portfolio very much, although Sharper Angle is closer to Markowitz in this case. The choice of the 20% tracking error is for ease of comparison only; this is a question of the *magnitude* of the weights vector, and it should be separated from the *direction* question.

7 Conclusion

We have suggested separating the mean-variance optimization into two steps, direction and magnitude. We argued that managers should monitor and constrain the angle between their alphas and positions in much the same way they monitor and constrain their positions in individual assets, sectors, and markets.

Standard optimized investing is not equipped to deal with the presence of estimation error, for both alphas and covariance. We have illustrated with a global bond spread trade, levered up by the optimizer because of its low relative risk, and subsequently "blown up" by a Black Swan event of the 2008 financial crisis. This problem could be much more insidious in a portfolio of hundreds or thousands of assets involving complicated constraints under rapidly changing market conditions. The manager may not be fully aware of all the complex low volatility spread trades implicitly levered up by the optimization "black box."

We have proposed a simple spherical "uncertainty cloud" model for the alphas, and used it in a robust optimization framework. Our robustly optimized portfolio is better positioned to reduce the leverage of the "noise" alphas, perhaps at the expense of a lower ex-ante Sharpe Ratio, and control risk in a regime-changing environment.

The geometric "angle" point of view is applicable only for a portfolio of similar assets. Allocating capital among assets with very different intrinsic leverage profiles is a very important problem for a fund-of-funds manager trying to optimize alphas of different natures. Thus, our proposed first step in the leverage control process is to put the optimization *input* alphas on the same inherent-leverage footing *before* throwing them into our proposed - or in fact any - optimization machine, lest it produces a "garbage in, *levered* garbage out" result. With inputs appropriately assembled, in the second step the robust optimization successfully controls the *direction* of the weights vector and thereby constrains the unintended leverage of smaller alphas. The third step is to define the overall leverage by adjusting the weights vector *magnitude*.

Further research directions include extending our results in the presence of position constraints, monitoring angles between subsets of alphas and weights, and generalizing to ellipsoidal “uncertainty clouds.”

Appendix: Degeneracy numbers

Recall that we defined the *minimax degeneracy* number of the matrix Σ as

$$\mathrm{mmd}(\Sigma) = \frac{\theta_{\max}\theta_{\min}}{(\theta_{\max}^2 + \theta_{\min}^2)/2} = 2\sqrt{\frac{\kappa}{(\kappa+1)^2}} \le 2\sqrt{\frac{1}{\kappa}}, \tag{16}$$

It could also be useful to consider another condition-like number associated to the matrix Σ, see Axelsson (1996)

$$\tilde{\kappa} = \frac{\mathrm{tr}\Sigma/N}{(\det\Sigma)^{1/N}} \le \kappa, \tag{17}$$

and define the *full degeneracy* number

$$\mathrm{dec}(\Sigma) = \sqrt{\frac{(\det\Sigma)}{(\mathrm{tr}\Sigma/N)^N}} = \sqrt{\left(\frac{1}{\tilde{\kappa}}\right)^N}. \tag{18}$$

If $\Sigma = DCD$ where $D = \mathrm{diag}(\sqrt{\Sigma_{11}}, \sqrt{\Sigma_{22}}, \ldots, \sqrt{\Sigma_{NN}})$ and C is the correlation matrix, then

$$\mathrm{dec}(\Sigma) = \mathrm{dec}(D^2)\mathrm{dec}(C). \tag{19}$$

This equation *does not hold* for the minimax degeneracy number $\mathrm{mmd}(\Sigma)$. The two degeneracy numbers are related by the following inequalities

$$2(\mathrm{dec}(\Sigma))^{1/N} \ge \mathrm{mmd}(\Sigma) \ge \mathrm{dec}(\Sigma). \tag{20}$$

Finally, in the case where α and w lie in a purely 2-dimensional eigen-subplane of Σ, we can obtain a better lower bound on the cosine, refining (12) to

$$\cos(\omega) \ge \mathrm{dec}_2\sqrt{1-\chi^2(1-\mathrm{dec}_2^2)} + \chi(1-\mathrm{dec}_2^2), \tag{21}$$

where dec_2 is the full degeneracy number of the restriction of Σ to that 2-plane. Thus both the parameter χ and the degeneracy number of Σ contribute toward keeping the angle ω sufficiently acute.

References

Axelsson, O. (1996) *Iterative Solution Methods*. Cambridge, UK: Cambridge Univ. Press.

Black, F., and Litterman, R. (1992) “Global Portfolio Optimization,” *Financial Analysts Journal* 48 (5), 28-43.

Ceria, S. and Stubbs, R. (2006) “Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction,” *Journal of Asset Management* 7 (2), 109-127.

Chopra, V.K. and Ziemba, W.T. (1993) “The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice.” *Journal of Portfolio Management* 19 (2), 6-11.

Fabozzi, F., Kolm, P., Pachamanova, D., and Focardi, S. (2007) *Robust Portfolio Optimization and Management* J. Wiley & Sons.

Goldfarb, D., and Iyengar, G. (2003) "Robust Portfolio Selection Problems." *Mathematics of Operations Research* vol 28, No 1, 1-38.

Jacobs, B., Levy, K., and Markowitz, H. M. (2006) "Trimability and Fast Optimization of Long-Short Portfolios." *Financial Analysts Journal* 62, 36-46.

Jorion, P. (1986) "Bayes-Stein estimation for portfolio analysis." *Journal of Financial and Quantitative Analysis* 21, 279-291.

Jorion, P. (1992) "Portfolio Optimization In Practice." *Financial Analysts Journal* 48 (1), 68-74.

Kritzman, M. (2006) "Are Optimizers Error Maximizers?" *Journal of Portfolio Management*, Summer 2006, 66-69.

Markowitz, H. M. (1952) "Portfolio Selection." *Journal of Finance* 7, 77-91.

Markowitz, H. M. (1959) *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons; 2nd ed., 1991. Cambridge, MA: Basil Blackwell.

Markowitz, H. M. (1987) *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. New York: Basil Blackwell.

Meucci, A. (2005) "Robust Bayesian Allocation." Available at SSRN: http://ssrn.com/abstract=681553.

Michaud, R. O. (1989) "The Markowitz Optimization Enigma: Is Optimized Optimal?" *Financial Analysts Journal* 45, 31-42.

Michaud, R. O. (1998) *Efficient Asset Management*, Boston, MA: Harvard Business School Press.

Scherer, B. (2005) "How Different is Robust Optimization Really?" Working paper.

Tütüncü, R. H., and Koenig, M. (2004) "Robust Asset Allocation." *Annals of Operations Research* 132, 157-187.

Endnotes

[1]This can always be done in the absence of effective nominal constraints; or in other words, for the *homogeneous* problem. We shall assume this throughout the paper. Practitioners may, however, introduce various asset-level and global constraints. This is an interesting variation of the problem, with possibly nontrivial consequences, but many of our ideas and results will still apply in that setting.

[2] Concerns about the leverage of the optimization inputs are especially relevant in multi-asset-class or multi-strategy capital allocation contexts.

[3]The literature on the covariance matrix estimation procedures is very ample. Without going into a discussion of these techniques, we assume here that a covariance matrix estimate has been given to us.

[4]In the classical mean-variance optimization framework, taking "risk" means surviving the volatility while waiting for the mean forecasts to be realized. Robust optimization deals with the "risk" of the mean forecasts to be wrong in the first place, and controls the damage for the "worst case" scenario. General robust frameworks also allow for uncertainty of the covariance matrix estimates. We will not address this possibility here.

[5] The spherical "uncertainty cloud" and the geometric "angle" interpretation works only for a portfolio of similar, or "homogeneous," assets. For anisotropic portfolios it would be more appropriate to consider ellipsoidal or other asymmetric uncertainty clouds.

[6] One of the most famous Bayesian shrinking methods is the Black-Litterman approach, which deals with the uncertainties in alpha and covariance estimates by changing the parameters of the normal distribution. However, Black-Litterman is conceptually different from the "worst case" robust methodology.

[7]It is possible to show that the solutions of these problems (so-called Second Order Cone Programming problems) always exist and it is not hard to find these solutions numerically.

[8]In certain specific robust problem settings, solutions reduce to elementary "shrinkage" and can be understood and argued outside of the robust framework, a fact appreciated in Scherer (2005). We stress that here, the dependence of t on Σ and α means that our solution is nontrivial.

[9]All performance charts are rescaled to 5% tracking error.

[10]We chose an example with exaggerated German/French trades for this illustration. In practice, a portfolio manager may use constraints, discretionary overrides, or other procedures in order to avoid such spread trades. However, the performance of our Manager B would still suffer in Q4 2008 had he/she used reasonably forgiving individual position and global constraints.