Robust Risk: Using Robust Methods to Improve Investment Performance

Roy Welsch

Massachusetts Institute of Technology

rwelsch@mit.edu

Interface Symposium
May 21-24, 2008
Durham, NC

Notation

N = number of risky assets

μ_i = expected return of the ith asset, μ = N-vector of expected returns

Σ = covariance matrix of the returns of N assets

T = number of time periods for estimating Σ and μ.

$\mathbf{r}_i$ = returns of the ith asset in the portfolio (a vector of length T)

w_i = weight of the ith asset in the portfolio, $\sum_{i=1}^{N} w_i = 1$

$\mathbf{R}_t$ = returns at time t (a vector of length N)

Note: w_i could be negative for short sales.

Mean-Variance Portfolio Optimization

Portfolio return:

$$r_p = w_1 r_1 + w_2 r_2 + \ldots + w_N r_N$$

Expected portfolio return

$$E(r_p) = \mathbf{w}'\mu$$

Variance of the portfolio return

$$\mathrm{Var}(r_p) = \mathbf{w}'\Sigma\mathbf{w}$$

Mean-variance portfolio optimization minimizes the variance of a portfolio return for a given level of expected return μ_p:

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

subject to $\mathbf{w}'\mu \geq \mu_p$, $\mathbf{w}'\mathbf{e} = 1$

where $\mathbf{e}$ is the $N \times 1$ column vector with all elements 1.

No constraint on the mean return gives the minimum variance portfolio (MVP).

Problems with Mean-Variance

1. Sensitive to inputs which are, in turn, subject to random errors in the estimation of the expected return and variance that are, usually, obtained from historical return data.
2. This sensitivity often leads to extreme portfolio weights and dramatic swings in weights with only minor changes in expected returns or the covariance matrix. This can lead to frequent rebalancing and excessive transaction costs.
3. For stable covariance estimation, we prefer long historical time series (the number of assets, N, far smaller than the number of time periods, T). However, old historical data may not reflect current market dynamics.
4. Underlying multivariate normal assumption (model risk) may not be right.

Some Solutions

1. Factor models (CAPM, etc.), Bayesian shrinkage, GARCH models.

2. Regularization (penalty) methods.

3. Robust estimation of the expected return and the covariance matrix.

4. Robust optimization.

5. Combinations of the above methods.

Contamination Models

Equivariant (MCD):

$$F = (1-\varepsilon)F_0 + \varepsilon H \qquad 0 < \varepsilon < \frac{1}{2}$$

Each row (time observation) either from F_0 or H. Implies either a bad day on the market (all stocks) or a high correlation among stocks. In fact, rarely true.

Pairwise:

Pairwise correlation permits a more flexible error model. Unusual market returns only explain a small part of observed outliers. Industrial factors and idiosyncratic risk specific to individual stocks or groups of stocks explain a majority of the outlying data.

Fast-MCD

The minimum covariance determinant (MCD) proposed by Rousseeuw in1985 looks for the covariance matrix of h data points ($T / 2 \leq h < T$) with the smallest determinant. The breakdown is $(T - h) / T$. The resulting covariance matrix is biased (and can be adjusted to be unbiased), but this multiplicative factor has no effect on portfolio weight allocation. MCD is not feasible for $N > 20$ in our situation. Fast-MCD proposed by Rousseeuw and Van Driessen (1999) makes large N feasible. MCD retains affine equivariance.

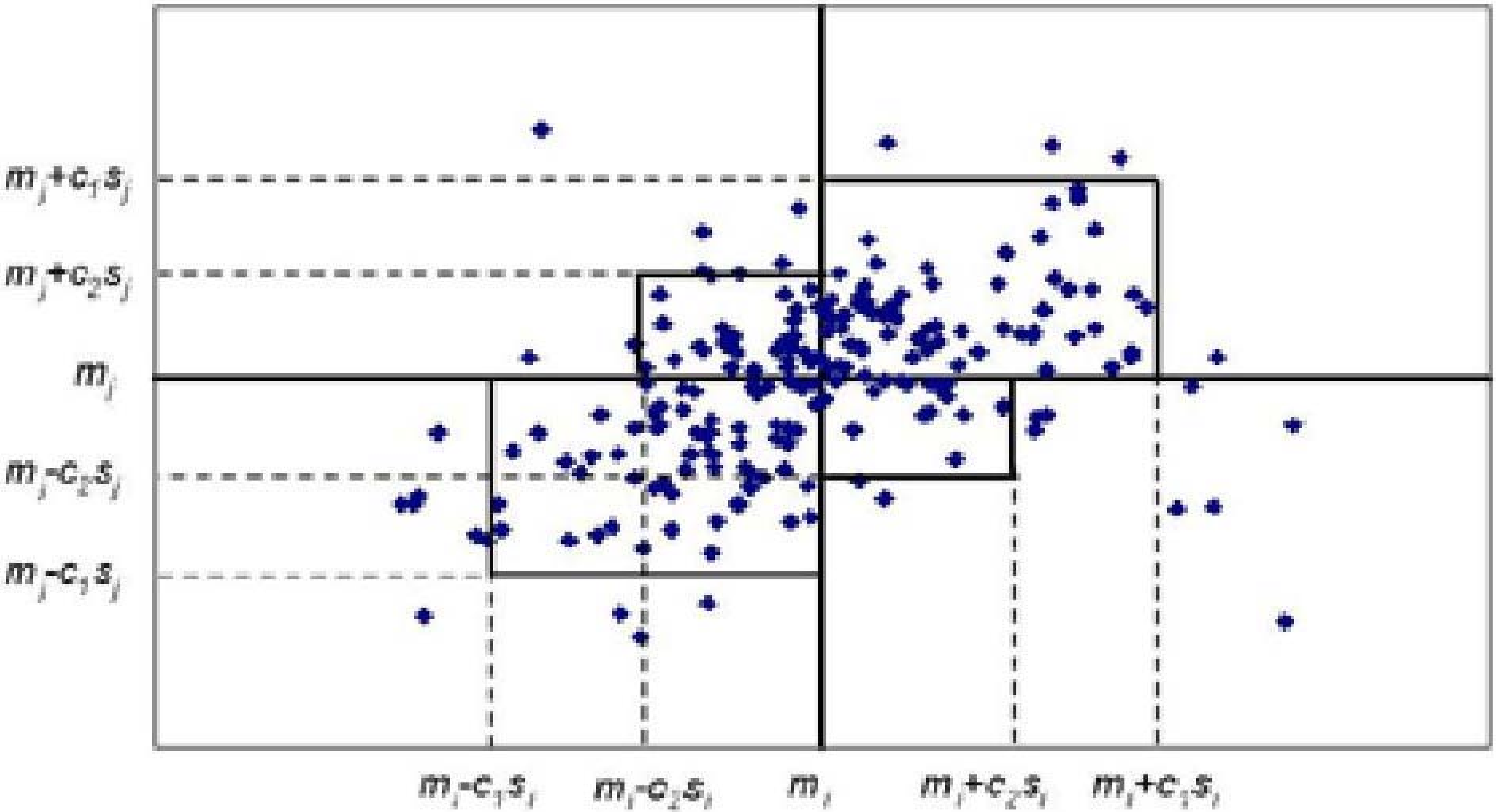
Pairwise Robust Covariance

If we do not demand affine equivariance, then fast robust pairwise covariance estimators are available.

Khan et.al. (2005) compared several approaches to robust pairwise covariance estimation while investigating ways to make least-angle regression (LARS) (Efron, et. al., 2003) robust. They found a two-step, two-dimensional Winsorization method to be effective and fast. We use a modified form of their idea with adjustment to insure a positive definite covariance matrix.

Fast 2-D Winsorization

Khan, et. al. (2005) start with univariate Huber Winsorization but use two tuning constants, c_1 and c_2. The constant c_1 (chosen to be 2) is used in the two quadrants with the most data (n_1) and the second constant $c_2 = n_2 / n_1$ with $n_2 = T - n_1$ is used in the remaining two quadrants. This pulls the Huber Winsorization boundary in where there is less data and a higher chance of data not following the ellipsoidal pattern for the main part of the data.



Adjusted Winsorization (for initial covariance) with $c_i = 2$, where s_i and s_j are estimated from adjusted MAD

Fast 2D Winsor

The classical correlation is computed on this Winsorized data and all pairwise correlations form the full initial correlation matrix which, if necessary, is made positive definite. Then one step of bivariate Winsorization is used and this new matrix is again made positive definite. We call this I2DW.

Many other ways to do this.

Robust Optimization

- Use uncertainty sets for μ and Σ since estimates for exact values are uncertain.
- Find weights, **w**, that minimize portfolio risk for the worst case values of $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ in their uncertainty sets
- For larger problems use a factor model

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

V = factor loadings
f = stochastic vector of m factors with $m << N$.
ε = stochastic vector of residual returns
The capital asset pricing model (CAPM) uses one factor – return of a value-weighted average of all assets in the market.

Uncertainty Sets

Uncertainty sets for **V** and Σ_ε, and μ in place of sets for Σ and μ greatly simplify the computations. Often assume

$$\mathbf{f} \sim N(\mathbf{0}, \mathbf{F}), \mathbf{F} \text{ known and stable}$$

$$\varepsilon \sim N(\mathbf{0}, \Sigma_\varepsilon)$$

$$\Rightarrow \mathbf{r} \sim N(\mu, \mathbf{V}' \mathbf{F} \mathbf{V} + \Sigma_\varepsilon)$$

V determined by regression using data for the factors. Uncertainty sets are based on these regressions and residuals.

Uncertainty Set Examples

$$S_{\boldsymbol{\varepsilon}} = \left\{\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} : \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = \mathrm{diag}(\boldsymbol{\sigma}), \sigma_i \in \left[\sigma_i^-, \sigma_i^+\right], i = 1, \ldots, N\right\}$$

$$S_{\boldsymbol{\mu}} = \left\{\boldsymbol{\mu} : \boldsymbol{\mu} = \hat{\boldsymbol{\mu}} + \mathbf{v}, \left|v_i\right| \le \gamma_i, i = 1, \ldots N\right\}$$

$$S_{\mathbf{V}} = \left\{\mathbf{V} : \mathbf{V} = \hat{\mathbf{V}} + \mathbf{W}, \left\|W_i\right\|_g \le \rho_i, i = 1, \ldots N\right\}$$

where the *g*-norm is also based on the factors obtained from historical data. Details may be found in Goldfarb and Iyengar (2003).

The Robust Portfolio Optimization Problem

The robust portfolio optimization problem becomes:

$$\min_{\mathbf{w}} \max_{\mathbf{V} \in S_V, \boldsymbol{\Sigma}_\varepsilon \in S_\varepsilon} \mathbf{w}'\boldsymbol{\Sigma}\mathbf{w}$$

$$\text{s.t.} \quad \min_{\boldsymbol{\mu} \in S_\mu} \mathbf{w}'\boldsymbol{\mu} \geq \mu_0 \qquad \mathbf{w}'\mathbf{e} = 1$$

This problem can be reformulated into a second-order cone problem that is easier to solve. From a practical point of view, the trick is to formulate useful and realistic uncertainty sets that, in fact, can be transformed into solvable optimization problems.

Uncertainty Set Issues

- Generally think of these sets as reflecting future uncertainty in investment returns
- Do we expect outliers and bad data?
- Do we want to use robust covariances on historical data to determine elliptical uncertainty sets when outliers may occur in the future? If not, then overly conservative?
- If so, then we are more or less back at square one needing to find robust covariances.
- May need to consider a number of scenarios for the future.
- Moral: We may still need robust estimation even in the presence of robust optimization!

Minimum Variance Portfolio

The classical minimum variance portfolio optimization problem,

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}'\Sigma\mathbf{w}$$

$$\text{s.t.} \quad \mathbf{w}'\mathbf{e} = 1$$

is equivalent to

$$\min_{\mathbf{w},v} \sum_{t=1}^{T}\left(\mathbf{R}_t'\mathbf{w} - v\right)^2 \text{ or } \min_{\tilde{\mathbf{w}}} \sum_{t=1}^{T}\left(\tilde{\mathbf{R}}_t'\tilde{\mathbf{w}} - 0\right)^2$$

if the intercept is included in $\tilde{\mathbf{R}}_t$ and $\tilde{\mathbf{w}}$.

We can then use robust regression and/or robust optimization.

Too Much Turnover?

One way to possibly reduce turnover and instability would be to penalize deviations from the market weights, **m**, and, at the same time, look for sparse solutions that do not invest any funds in some securities. The LASSO (Tibshirani, 1996) does exactly this.

Penalization

To implement this we solve (Lauprête, 2001)

$$\arg\min_{\mathbf{w},v} \frac{1}{T}\sum_{t=1}^{T}(\mathbf{R}_t'\mathbf{w}-v)^2+\lambda\sum_{i=1}^{N}|w_i-m_i|$$

and use 5-fold cross-validation to find λ based on prediction error for the out-of-sample data. The recently developed LARS (least-angle regression) algorithm (Efron, et. al. 2004) greatly speeds up computations for the Lasso since solutions for all λ can be found in about the same time as one least-squares regression. This removes the need for a (non-specific) grid search on λ. Here m_i is some benchmark portfolio.

Financial Performance Measures

SD: the sample standard deviation of ex-post returns.

Sharpe ratio:

(average annual return – r_f) / annual SD

where r_f is the risk free rate.

T**urnover**: Asset turnover, defined as the mean of the absolute weight changes $\sum_{i=1}^{N} \left| w_{i,t} - w_{i,t-1} \right|$ for all updates.

Historical Data

We tested robust optimization and robust estimation on a set of real data that contains daily returns from 15 midcap stocks from 1/2/2003 to 12/29/2006.

Rebalance as follows: Estimate sector weights for the next week using most recent $T = 100$ daily returns, rebalance every 5 trading days and roll training window forward 5 days.

Methods

The labels correspond to the following:

- $1/N$: Portfolio obtained from equal weights
- MVP: Minimum variance solution
- I2DW: Iterated bivariate Winsorization
- MCD: Robust equivariant
- RO: Robust optimization
- FAC: Factor method with 5 factors

MV Numerical Results

907 Days	1 / *N*	V	I2DW	MCD	RO	FAC
Mean	0.378	0.422	**0.435**	0.411	0.354	0.415
SD	0.197	0.188	0.191	0.189	**0.187**	0.188
Sharpe	1.663	1.984	**2.019**	1.905	1.625	1.940
Wealth	3.061	3.419	**3.528**	3.323	2.882	3.359
Turn	**0**	0.270	0.347	0.377	0.516	0.379

Next Steps

1. Robust estimators for uncertainty sets

2. Penalized robust methods

3. Parameter cross-validation

4. Faster algorithms

References

DeMiguel, V. and Nogales, F.J. (2008). Portfolio selection with robust estimation of risk, to appear in *Operations Research*.

Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004). Least angle regression, Annals of Statistics, **32**, 2, 407–499.

Goldfarb, D., Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research* , **28**(1), 1–38.

Khan, J., Van Aelst, S. and Zamar, R. (2005) Robust linear model selection based on Least Angle Regression, Technical Report, Department of Statistics, University of British Columbia.

Laupréte, G.J., Portfolio risk minimization under departures from normality. MIT PhD Thesis, 2001.

Lauprete, G., Samarov, A. and Welsch, R. (2002). Robust portfolio optimization, Metrika, 55, 139–149.

Maronna, R.A., Martin, R.D. and Yohai, V.J. (2006). *Robust Statistics: Theory and Methods*, Wiley, New Jersey.

Maronna, R. and Zamar, R. (2002) Robust estimates of location and dispersion for high-dimensional datasets, *Technometrics*, **44**(4), 307–317.

Menjoge, R., Nguyen, D., and Welsch, R. (2007) Robust estimation or robust optimization for portfolio selection: Which should you use? *Proceedings of the 56th Session of the International Statistical Institute*.

Rousseeuw, P.J. and Van Driessen, K. (1999) A fast algorithm for the minimum covariance determinant estimator, *Technometrics*, **41**(3), 212–223.

Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society, Series B*, **5**, 267–288.

Welsch, R. and Zhou, X. (2007) Application of Robust Statistics to Asset Allocation Models, *REVSTAT*, **5**, (1), 97–114.

Zhou, X. (2006) Application of Robust Statistics to Asset Allocation Models, MIT, M.S. Thesis.