Portfolio Optimization Without Forecasts

Robert Almgren
University of Toronto

Joint work with Neil Chriss, SAC Capital
Finance has good math problems

- Important—come from real challenges.
- Messy—math should be simple but good.
- Much more than option pricing.
Earlier example

Joint work with Neil Chriss 1997
Problem from Morgan Stanley program trading

Liquidate large investment, minimizing market impact costs and volatility risk.

Careful modeling assumptions:
Calculus of variations → optimal trajectories
Current interesting example

Motivated by day-to-day problems

1. Classic portfolio optimization

2. Difficulties encountered in real life

3. New notion of optimality
   (a) Partial ordering
   (b) Full ordering

4. Empirical tests

Preprint forthcoming.
1. Mean-Variance Optimization

(Markowitz 1952)
Portfolio manager has $n$ assets
(thousands of US stocks)
Must allocate wealth in order to

- Maximise return (growth), and
- Minimise risk

Use known parameters of asset price motion.
Brownian motion

Prices $S_i(t)$ give one-period returns ($\xi_i \sim \mathcal{N}(0, 1)$)

$$
\rho_i \equiv \frac{S_i(T) - S_i(0)}{S_i(0)} = r_i T + \sigma_i \sqrt{T} \xi_i,
$$

$$
\mathbb{E}(\rho_i) = r_i T \quad \mathbb{E}((\rho_i - r_i T)(\rho_j - r_j T)) = V_{ij} T
$$

$$
R = (r_1, \ldots, r_n) = \text{expected returns}
$$

$$
V = \text{covariance matrix}
$$

$R$ predicted from analysis and forecast

$V$ from historical data (assume stable)
Portfolio investment

\( w_i = \) fraction of wealth invested in \( i \)th asset

(Can borrow if desired)

\[
\text{Portfolio return} = w^T R \quad \text{(maximise)}
\]

\[
\text{Return variance} = w^T V w \quad \text{(minimise)}
\]
Two ways to include risk

1. Quadratic utility, \( \lambda = \) risk-aversion parameter

\[
\max_{w \in \mathbb{R}^n} \left( w^T R - \lambda w^T V w \right)
\]

2. Limited “risk budget”

\[
\max_{w^T V w \leq \sigma^2} w^T R
\]

Solution: \( w \propto V^{-1} R \), suitably normalized
Choose high-return assets, relative to volatility.
2. Difficulties

1. \( w \) is very sensitive to \( R \), and \( R \) is imperfectly known
   \( \implies \) robust optimization techniques

2. May have only ordered list of assets, not quantitative values for \( r_i \).
   \( \implies \) subject of this work
Simple experiment:

- Use expected returns and covariance from Idzorek (2002) for Dow 30
- Randomly generate 10,000 expected return estimate vectors from a normal distribution with mean equal to the expected return and std equal to 0.1% of the std of return of the corresponding asset
- Run 10,000 traditional MVO and record the weights of the resulting portfolios – Use a fixed risk aversion coefficient

![Figure 1: Range of Expected Returns used in Mean-Variance Optimizations](image1)

![Figure 2: Range of Asset Weights in Mean-Variance Optimal Portfolios](image2)
Why lack of quantitative values

- Estimates come from qualitative value somehow correlated with $R$
  - Reversal (Campbell/Grossman/Wang)
  - Momentum (Asness)
  - Book/market and size effects (Fama/French)
  but precise relationship to $R$ not known
- Come from “secret” source (buy a list)
- Lack of confidence in numbers
  (alternative to robust regression)
Current practice

1. “Buy the top, sell the bottom”
   • Long position in first asset (or first 10%)
     Short the last asset (or last 10%)
   • Linear weighting

2. Invent returns from ordering, use MVO.
   Ad hoc: linear? inverse cumulative normal?
3. Portfolio Comparison

$w_1 \succeq w_2$ means “we prefer $w_1$ over $w_2$$$
MVO: w_1 \succeq w_2 \iff w_1^T R \succeq w_2^T R$

Want “most preferable” $w$ within constraints, e.g.

- Risk limit $w^T V w \leq \sigma^2$
- Market-neutrality $w^T e = 0, e = (1, \ldots, 1)^T$
- Total investment $|w_1| + \cdots + |w_n| \leq W$
- Position constraints, \textit{etc.}$\$

Maximise scalar function $f(w) = w^T R$

What to maximise when don’t have $R$?
Two steps to construct ordering

(a) $w_1 \succeq w_2$ means “$w_1$ is at least as good as $w_2$ in every scenario consistent with our info.”
Partial ordering: large set of extrema.

(b) $w_1 \preceq w_2$ if $w_1$ better “more often”
Full ordering: unique best portfolio.
**Cone definition**

Standard case: have expected return $R$

$$\mathcal{R} = \{ \lambda R \mid \lambda > 0 \}$$

$$w_1 \succeq w_2 \iff w_1^T r \geq w_2^T r \text{ for all } r \in \mathcal{R}.$$ 

*Magnitude* of return $R$ does not matter only *direction*.

Maximise exposure to “good directions”

Don’t care about orthogonal directions
Inequality beliefs

$m$ inequality statements

$D_i^T r \geq 0, \ i = 1, \ldots, m$

$Dr \geq 0$ in $\mathbb{R}^m$

Simple sort: $r_1 \geq \cdots \geq r_n$, $m = n - 1$

$$D = \begin{pmatrix} D_1^T \\ \vdots \\ D_m^T \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ & \ddots & \ddots \\ & & 1 & -1 \end{pmatrix}$$
More general beliefs

$k$ sectors

\[ r_1 \geq \cdots \geq r_{n_1} \]
\[ r_{n_1 + 1} \geq \cdots \geq r_{n_2} \]
\[ \vdots \]
\[ r_{n_{k-1} + 1} \geq \cdots \geq r_n. \]

No information (no opinion) about relative ordering between sectors.

\[ Dr \succeq 0, \ D \text{ is } (n - k) \times n. \]
Consistent cone

\[ \mathcal{R} = \{ r \in \mathbb{R}^n \mid Dr \geq 0 \} \]

The set of expected returns that are \textit{consistent} with our beliefs.
We have no other knowledge or beliefs about \( r \).
$D$ is size $m \times n$, $m \leq n$

($n$ is number of assets, $m$ constraints)

Augment $D$ with “dummy” rows $D_{m+1}, \ldots, D_n$

orthogonal to “real” rows $D_1, \ldots, D_m$.

$\tilde{D}$ is $n \times n$ nonsingular

Inverse $\tilde{E}$: $\tilde{D}\tilde{E} = \tilde{E}\tilde{D} = I_n$.

$E =$ first $m$ columns of $\tilde{E}$
Change of basis

\[ w^T r = w^T (\tilde{E} \tilde{D}) r = (\tilde{E}^T w)^T (\tilde{D} r) = v^T s \]

\[ r \in \mathcal{R} \iff s \in S \]

\[ S = \{ s \in \mathbb{R}^n \mid s_1, \ldots, s_m \geq 0 \} \]

Definition:

\[ w_1 \geq w_2 \iff \tilde{E}^T (w_1 - w_2) \in S \]

\[ \iff w_1 - w_2 = \tilde{D}^T v, \quad v \in S \]

Positive exposure to “difference assets”

Don’t care about exposure to other directions.
Consistent returns

\[ r_1 \geq r_2 \]

Portfolios \( w \geq w^0 \)

\[ w - w^0 = E_1 \nu, \nu_1 \geq 0 \]
Portfolios $w > w^0$

$w - w^0 = E_1 v$, $v_1 > 0$

**Strict preference**

$w_1 > w_2 \iff w_1 \succeq w_2, w_2 \not\succeq w_1$
Look for extremal portfolios over allowable set of weights

\[ \mathcal{M} = \{ w \in \mathbb{R}^n \text{ that satisfy constraints} \} \]

Risk ellipsoid: \[ \{ w^T V w \leq \sigma^2 \} \]

Market neutral: \[ \{ w^T V w \leq \sigma^2 \text{ and } w^T e = 0 \} \]

Total investment: \[ \{ |w_1| + \cdots + |w_n| \leq W \} \]
2 (a). Extremal portfolios

\( w^* \in \mathcal{M} \) is “extremal” if there is no \( w \in \mathcal{M} \) with \( w > w^* \).

Main result:

Extremal \iff Condition on normal
(to some supporting hyperplane)
\[ \text{Pos}(E) = \text{positive cone of columns of } E \]

**Theorem** Suppose \( \mathcal{M} \) has a normal \( b \) at \( w^* \) with \( b \in \text{Pos}(E) \). If \( b \) is a strict normal, or if \( b \) is in the strict positive cone, then \( w^* \) is maximal in \( \mathcal{M} \).

**Theorem** Suppose that \( \mathcal{M} \) is convex. If \( w^* \) is maximal in \( \mathcal{M} \), then there is a normal \( b \) to \( \mathcal{M} \) at \( w^* \) so that \( b \in \text{Pos}(E) \).

**Proof** Separating Hyperplane Theorem.
Total Investment Constraint

\[ |w_1| + \cdots + |w_n| \leq W \]

\[ w^* = (w_1, 0, \ldots, 0, w_n) \]

\[ w_1 \geq 0, \ w_n \leq 0, \ w_1 - w_n = W \]

1-dimensional line segment

Long top, short bottom, in arbitrary amounts

Multi-sector: long/short in each sector
Risk Ellipsoid

\[ w^T V w \leq \sigma^2 \]

Unique normal to \( \mathcal{M} \) at \( w \) is \( b = Vw \)

\[ b = Ex \iff w = V^{-1}Ex, \quad x \geq 0 \text{ in } \mathbb{R}^m \]

Single sort \( m = n - 1 \):

(n - 2)-dimensional patch on \((n - 1)\)-dimensional ellipsoid.

Each point in region is extremal.
(not dominated by any other)
Market-Neutral on Risk Ellipsoid

\( \mathcal{M} \) is “equator,” dimension \( n - 2 \).

Normals are one-parameter family

\[
\mathcal{B} = \left\{ \alpha Vw + \beta D_n \mid \alpha \geq 0, \beta \in \mathbb{R} \right\}
\]

\[
w = V^{-1}Ex + \gamma V^{-1}D_n, \quad \gamma = -\frac{D_n^T V^{-1}Ex}{D_n^T V^{-1}D_n},
\]

Single sort \( m = n - 1 \):

\( (n - 2) \)-dimensional patch on \( (n - 1) \)-dimensional ellipsoid.
3 assets
How to choose single optimum
One idea:
Take “central point” $x = (1, \ldots, 1)^T$
Equivalent to linear returns.

Less ad hoc?
3 (b). Optimal portfolios

\( w_1, w_2 \) both extremal:
Each has higher expected return than the other for \textit{some} expected return vector.

Refine the definition.
Idea: \( w_1 > w_2 \) if \( w_1 \) is \textit{more often} better.

Need to define “more often”
New definition:

\[ w_1 \geq w_2 \iff \text{meas}\{ r \in \mathcal{R} \mid w_1^T r \geq w_2^T r \} \geq \text{meas}\{ r \in \mathcal{R} \mid w_2^T r \geq w_1^T r \} \]

Includes previous definition (meas\{ \cdot \} = 0)

What measure to use?
Need \( \text{meas}(\mathcal{R}) < \infty \)
Expected return \( r \in \mathcal{R} \)
No other information

Assume \textit{radially symmetric} distribution within \( \mathcal{R} \).
(Assert Euclidean rotation is natural symmetry)

Example: Gaussian of radius \( \rho \) (any value of \( \rho \))

\[
P(r) \, dr \propto \exp\left(\frac{|r|^2}{2\rho^2}\right) 1_{\mathcal{R}} \, dr
\]
Geometric description

Level sets of relation \( \succeq \)

\[
\begin{align*}
  w_1 \simeq w_2 & \iff \text{meas}\{ r \in \mathcal{R} \mid (w_1 - w_2)^T r \geq 0 \} \\
  & = \text{meas}\{ r \in \mathcal{R} \mid (w_1 - w_2)^T r \leq 0 \}
\end{align*}
\]

\( \iff \) Plane with normal \( w_1 - w_2 \) bisects \( \mathcal{R} \)
passes through centroid \( c \) of \( \mathcal{R} \)
\( \iff (w_1 - w_2)^T c = 0 \)
Characterization of ordering

\[ w_1 \geq w_2 \iff w_1^T c \geq w_2^T c \]

Centroid vector \( c \) is effective return!
Is extremal by previous definition.

Compute optimal elements just as for classic MVO, including constraints.
Example: Risk limit \( w \propto V^{-1} c \)
Computing centroid

Two methods:

1. Monte Carlo: generate Gaussian samples. Sorting brings into $R$ by reflection.

2. Order statistics

   Blom approximation (1958):

   $c_{j,n} \approx N^{-1} \left( \frac{n + 1 - j - \alpha}{n - 2\alpha + 1} \right)$

   use numerical fit for $\alpha(n)$. 
Centroid compared to linear
—weights extremes more highly

Asset number

Centroid component
Same construction for sectors
Extensions

- Multiple incompatible sorts
  *e.g.*, reversals *and* book/market
  ⇒ Use combined centroid

- Uncertainty within ranges
  *e.g.*, divide into deciles
  ⇒ Average centroid within ranges

- Transaction cost penalties
4. Empirical tests

- Test whether makes sense
- Clarify role of imperfect information
Real data

- 100 large-cap stocks
- 3 years daily data
- One-week reversal strategy
  (sort by negative return in previous week)
Five strategies

1. Long top 1 stock, short bottom 1
2. Long top 10 stocks, short bottom 10
3. Linear weighting
4. Risk-adjusted linear: $V^{-1} \times \text{linear}$
5. Our optimal: $V^{-1} \times \text{centroid}$
The graph shows the cumulative return over time for different categories and risk measures. The x-axis represents time in days, ranging from 0 to 800. The y-axis shows the cumulative return in percentages, ranging from -50% to 300%. The categories include Top 1 - bottom 1, Top 10 - bottom 10, Linear, Risk linear, and Risk centroid. The graph indicates trends and comparisons across these categories.
Quantitative measurement

Mean return: week-to-week
Std Dev: of weekly returns, not daily
(close to 1 by risk normalization,
greater than 1 from rebalancing)
Sharpe ratio: Mean / Std Dev
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Sharpe</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1113</td>
<td>1.300</td>
<td>1.354</td>
</tr>
<tr>
<td>2</td>
<td>0.0496</td>
<td>1.117</td>
<td>0.703</td>
</tr>
<tr>
<td>3</td>
<td>0.0480</td>
<td>1.136</td>
<td>0.668</td>
</tr>
<tr>
<td>4</td>
<td>0.3456</td>
<td>1.157</td>
<td>4.724</td>
</tr>
<tr>
<td>5</td>
<td>0.3691</td>
<td>1.181</td>
<td>4.942</td>
</tr>
</tbody>
</table>
Simulated data

- 100 assets
- Return distribution linear or normal
- Volatilities distributed by Zipf’s law (like real data) max/min from 1 to 20
- Introduce permutations to degrade info (length = minimum number of pair reversals)
- Strategies normalized to unit risk

Look at \[rac{\text{Sharp ratio of optimal portfolio}}{\text{Sharpe ratio of linear portfolio}}\]
Permutation degrades improvement

Volatility Ratio = 1

Volatility Ratio = 20

Improvement in Sharpe Ratio

Permutation Length / Number of Assets
Patent Pending

“On the use of the centroid for forming optimal investment portfolios...”

License to Axioma, Inc.
Summary

- There are good math problems in finance
- Novel way to construct optimal portfolios
- Empirical tests are very good
- Mathematically interesting, and very practical