Portfolio Optimization: The Quest for Useful Mathematics

Stanley R. Pliska

University of Illinois at Chicago

Rothschild Visiting Professor
Isaac Newton Institute, Cambridge

www.uic.edu/∼srpliska

SOME OBJECTIVES

- Talk about the history of portfolio optimization
- Talk about some of the important, main theoretical ideas
- Talk about portfolio management in practice
- Suggest why optimization techniques are rarely used in practice, at least for portfolios of individual stocks
- Suggest that there are nevertheless opportunities for mathematical optimization to be used in practice
- Suggest some promising directions for future research

(This is largely a personal quest!)

Markowitz's Model

- Single period
- m securities, each with return R_i
- $r_i = E[R_i], \sigma_{ij} = Cov(R_i, R_j)$
- π_i is initial portion of wealth invested in security i
- The investor's objective is to find $(\pi_1, ..., \pi_m)$ so as to

Minimize $\sum_{i,j} \sigma_{ij}\pi_i\pi_j$ (variance = risk)

Subject to:

- $\sum_i r_i\pi_i = R$ (target portfolio mean return)
- $\sum_i \pi_i = 1$ (budget constraint)
- $\pi_i \geq 0, i = 1, ..., m$ (optional short sales constraints, etc.)

Example 1: Eleven Dow Jones Industrials

quarterly data, 1992-1995, $R = 27\%$

company	r_i (%)	r_{capm} (%)	π unrestricted (%)	$\pi \geq 0$ (%)
1	13.4	5.1	18.9	0
2	25.9	12.9	5.4	7.3
3	17.9	8.7	-16.8	0
4	29.2	13.5	1.2	30.0
5	-2.9	18.4	-4.6	0
6	21.2	6.9	10.4	12.6
7	30.1	1.0	21.3	38.6
8	20.0	-2.0	19.2	11.6
9	21.1	9.6	50.4	2.9
10	14.7	7.9	8.0	0
11	-10.0	16.9	-13.4	0

Efficient Frontiers (unrestricted and non-negative weights)

portfolio σ (%)	target R (%)	max weight (%)	min weight (%)
7.4	11.5	47	-21
7.8	15	40	-14
8.5	17	37	-11
9.3	19	33	-8
10.3	21	30	-5
13.9	27	50	-17
9.9	17.7	39	0
10.0	19	36	0
10.7	21	27	0
12.7	23	26	0
15.6	25	31	0
19.6	27	39	0

Some Problems

- Typical optimal weights have too large a range.
- The mean return of a security changes due to time and other factors (so do covariances).
- Even if the mean (one-year, say) return is constant over time, it takes decades to get an accurate estimate.
- The optimal weights are sensitive to mean return estimates.
- The optimal weights usually produce disappointing results when used on a real-time, prospective basis.
- Besides, portfolio managers want to use their stock picking "skills."

Fisher Black to the Rescue

- If you add a riskless security having return R_0, then in (σ, R)-space the new efficient frontier is a straight line passing through $(0, R_0)$ and is tangent to the efficient frontier of just the risky securities.
- Economists say that in economic equilibrium this point of tangency must correspond to the actual market portfolio.
- So Black's idea is take the actual market capitalizations as weights, and then solve for the r_i's such that the optimal weights equal the market weights.
- A table shows these computed values (denoted r_{capm}), assuming the market's mean return is 10% and the market weights are equal.
- You compare these computed r_i's with your estimates of them, and then you invest accordingly.
- Apparently many portfolio managers use some version of this approach (e.g., tracking a market index with perturbations).

Example 1: CAPM Implied Mean Returns

r_{capm} assumes equal market caps

company	r_i (%)	r_{capm} (%)	π unrestricted (%)	$\pi \geq 0$ (%)
1	13.4	5.1	18.9	0
2	25.9	12.9	5.4	7.3
3	17.9	8.7	-16.8	0
4	29.2	13.5	1.2	30.0
5	-2.9	18.4	-4.6	0
6	21.2	6.9	10.4	12.6
7	30.1	1.0	21.3	38.6
8	20.0	-2.0	19.2	11.6
9	21.1	9.6	50.4	2.9
10	14.7	7.9	8.0	0
11	-10.0	16.9	-13.4	0

Example 2: Stan's Retirement Fund (TIAA-CREF)

- Daily data: 1/1/92-12/31/2000 for estimation; 1/1/2001-4/8/2005 for evaluation
- Four funds: bank account, stock, social choice, and bond

fund:	bank account	stock fund	social choice	bond fund
mean ann. ret. (%)	4.54	12.40	11.41	6.50
sigma (annual, %)	0.07	12.59	9.19	3.82
stock correlation (%)	1.3			
so ch correlation (%)	2.8	94.6		
bond correlation (%)	6.8	17.8	36.9	

Efficient Frontiers with Unrestricted and Non-negative Weights

sigma %	mean return %	bank %	stock %	soc ch %	bond %
0.07	4.54	100	0	0	0
1.82	6	70	-9	29	10
4.31	8	29	-22	68	25
6.81	10	-12	-34	107	39
9.30	12	-54	-46	146	54
11.79	14	-95	-58	185	68
0.07	4.54	1	0	0	0
1.85	6	67	0	16	17
4.38	8	21	0	39	40
7.02	10	0	0	71	29
11.08	12	0	59	41	0

A Real Test: Simulation of Real-Time Investment

- Buy-and-hold from 1/1/2001 to 4/8/2005
- Three strategies: 100% stock fund, 100% social choice fund, 59-41% division of initial investment

strategy	mean ann ret %	sigma %
100% stock	0.02	16.93
100% soc ch	2.36	10.62
59-41% split	1.01	14.01

Example 3: Four U.S. Stock Indexes

- Daily data
- 6/1/95 - 12/31/2000 used for parameter estimation
- 1/1/2001 - 12/31/2004 used for evaluation

	Index A	Index B	Index C	Index D
mean ann ret %	18.45	17.33	9.76	13.45
sigma %	21.99	15.39	24.42	12.61
B correlation %	77			
C correlation %	82	57		
D correlation %	73	71	86	

Efficient Frontiers with Unrestricted and Non-negative Weights

mean ret %	sigma %	Index A %	Index B %	Index C %	Index D %
15.88	10.37	11	-3	-53	145
18	10.93	38	-8	-78	148
20	12.35	64	-13	-103	152
22	14.39	91	-18	-127	154
24	16.81	117	-23	-151	157
14.12	12.46	0	17	0	83
16	13.54	0	66	0	34
18	18.32	60	40	0	0

The Test: Simulated Investment 1/1/2001 - 12/31/2004

Five Strategies:

- Buy-and-hold each index
- Buy-and-hold with initial %'s (64,-13,-103,152)

Strategy:	Index A	Index B	Index C	Index D	64,-13,-103,152
mean ann return %	-5.84	4.73	1.80	14.74	16.99
sigma %	23.79	18.14	25.53	19.39	17.58

Now We Are Getting Somewhere!

Final Remarks About Single Period Markowitz

- Useful in some cases, but...
- Often of dubious value, especially for portfolios of individual stocks, because...
- Has difficulty coping with statistical estimates of mean returns.
- Works only with a single, fixed planning horizon.
- Ignores utility functions (which economists are fond of).
- Ignores dynamic trading strategies and thus can produce suboptimal trading strategies.

Multi-Period Optimal Portfolio Models

- Samuelson was the first to tackle multi-period portfolio optimization.
- He did this by applying dynamic programming in discrete time.
- Merton (Samuelson's student) was the first to apply dynamic programming in a continuous time setting.

The Continuous Time Optimal Portfolio Problem

The market has two securities:

the **bank account** with price process

$$B_t = \exp(\int_0^t r_s ds),$$

where r_t is the (possibly random) **riskless short interest rate**, and

a **risky asset** with price process

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t),$$

where W is a **Brownian motion**.

Trading Strategies and the Wealth Process

The **trading strategy** is an adapted process π_t describing the time-t proportion of wealth held in the risky asset.

The **wealth process** X_t (the time-t value of the portfolio) is then given by

$$dX_t = (1 - \pi_t)X_t r_t dt + \pi_t X_t(\mu_t dt + \sigma_t dW_t)$$

$$= r_t X_t dt + \sigma_t \pi_t X_t(\theta_t dt + dW_t),$$

where $\theta_t \equiv (\mu_t - r_t)/\sigma_t$ denotes the so-called **risk premium**.

The Optimization Problem

Given a **utility function** U (a concave, strictly increasing, real-valued function) and initial wealth $X_0 = x$, the investor seeks to choose an admissible trading strategy π so as to **maximize the expected utility of wealth** at a fixed planning horizon $T < \infty$.

The investor therefore needs to solve for the **value function**:

$$V(t, x) \equiv \sup_{\pi} E[U(X_T) | X_t = x]$$

The Dynamic Programming (Stochastic Control) Approach

Under some regularity assumptions, the value function will satisfy the Hamilton-Jacobi-Bellman equation

$$V_t(t,x) + \sup_{\pi} \Big(x(r_t + \pi\sigma_t\theta_t)V_x(t,x) + \frac{1}{2}\sigma^2 x^2 \pi^2 V_{xx}(t,x) \Big) = 0$$

$$V(T,x) = U(x)$$

From the solution one can deduce the maximizing π, i.e., the optimal trading strategy π^*.

Unfortunately, this PDE is difficult to solve; solutions are known for only a few cases.

Two Solvable Cases When Coefficients are Deterministic

Case 1: Log Utility ($U(x) = \ln x$)

$$\pi_t^* = \frac{\mu_t - r_t}{\sigma_t}$$

Case 2: CRRA Utility ($U(x) = \frac{x^\alpha}{\alpha}$ with $\alpha < 1, \alpha \neq 0$)

$$\pi_t^* = \frac{\mu_t - r_t}{(1-\alpha)\sigma_t}$$

Generalization to the Case of m Risky Securities

$\Sigma\Sigma'$ is the **covariance matrix**

μ is now a (column) vector of **appreciation rates**

1 denotes a column vector of 1's

Case 1: Log Utility

$$\pi_t^* = \Sigma^{-1}(\mu_t - r_t \mathbf{1})$$

Case 2: CRRA Utility

$$\pi_t^* = \frac{1}{1-\alpha}\Sigma^{-1}(\mu_t - r_t \mathbf{1})$$

Example 2 (Continued): Another Backtest

- Simulated trading from 1/1/2001 to 4/8/2005
- Four strategies: three from before plus one with continuous trading

strategy	mean ann ret %	sigma %
100% stock	0.02	16.93
100% soc ch	2.36	10.62
59-41% buy & hold	1.01	14.01
Merton $\alpha = .5$	1.95	2.36

Example 3 (Continued): Still Another Backtest

- Trade from 1/1/2001 - 12/31/2004
- Five earlier buy-and-hold strategies
- Continuous trading strategy with $\alpha = 5/6$ and $r = 4\%$ (constant)

Strategy:	Index A	Index B	Index C	Index D	Mix	Merton
mean ret %	-5.84	4.73	1.80	14.74	16.99	16.39
sigma %	23.79	18.14	25.53	19.39	17.58	14.47

Remarks About Merton's Optimal Portfolio Model

- Seems to be useful for indexes and mutual funds
- Seems to work better than Markowitz's buy-and-hold (although transaction costs due to frequent trading are ignored)
- But Merton's model not so good for portfolios of individual stocks, largely due to the statistical difficulties that plagued us before.

A Risk Neutral, Martingale Approach for Solving Merton's Optimal Portfolio Problem

The Idea:

First we find the **optimal attainable wealth**, that is, the random variable describing the terminal value of the portfolio under the optimal strategy.

Then we find the trading strategy that corresponds to this wealth, just like finding the trading strategy which **replicates the payoff of a European option**.

This approach often has computational advantages over the dynamic programming, PDE approach first introduced by Merton.

The Challenge: Identifying the Set of Attainable Wealths

If the securities market model is **complete**, that is if the payoff of every European option can be replicated by some trading strategy, then the random variable Z is an attainable wealth if and only if

$$x = E^Q[Z/B_T],$$

where x is the initial wealth, B_T is the time-T value of the bank account, and the expectation is respect to the **risk neutral probability measure** Q.

Hence for Step 1 we want to solve:

$$\text{maximize} \quad E[U(Z)]$$

$$\text{subject to:} \quad E^Q[Z/B_T] = x$$

Solving the Convex Optimization Problem

Introducing the **Lagrange multiplier** λ leads to the Lagrangian

$$E[U(Z)] - \lambda(E[LZ/B_T] - x),$$

where L is the appropriate **Radon-Nikodym derivative**.

The optimal attainable wealth then is

$$Z^{opt} = I(\lambda^* L/B_T),$$

where I is the inverse of the marginal utility function U' and where the scalar λ^* is chosen to satisfy

$$x = E[LI(\lambda^* L/B_T)/B_T].$$

Common utility functions lead to explicit formulas for Z^{opt}.

Application of the Risk Neutral Approach: Continuous Time Markowitz

- Surprisingly, the continuous time version of the mean-variance portfolio problem has been solved only relatively recently.
- Most of this research is by X.Y.Zhou and his colleagues.
- In "Continuous-Time Mean-Variance Portfolio Selection with Bankruptcy Prohibition" (*Mathematical Finance*, January 2005) Bielecki, Jin, XYZ and I studied the variation where, in addition to the usual target mean return constraint, there is a constraint that the portfolio's value must remain non-negative.
- We used the risk neutral approach to solve this problem.

Emerging from this Study: Markowitz Options

Suppose the portfolio involves the bank account and only one risky security, say a stock market index.

The optimal attainable wealth resembles a call option:

- Below the “strike” the payoff is zero
- Above the strike the payoff is a strictly concave, increasing function.

Markowitz options can be sold together with zero coupon bonds.

Is there a market for persons preparing for retirement?

Merton's Intertemporal Capital Pricing Model (ICAPM)

The risky assets now have price dynamics in the form

$$dS_i(t) = S_i(t)(\mu_i(Y_t)dt + \sigma_i(Y_t)dW_t),$$

where $\sigma_i(\cdot)$ is now a row vector, W_t is now a column vector of Brownian motions, and Y is a **factor process** governed by a general SDE:

$$dY_t = C_t dt + \Lambda_t dW_t$$

in which C and Λ are suitable (possibly random) coefficients.

The investor still wants to maximize the expected utility of terminal wealth X_T, only now the asset appreciation rates and volatilities depend explicitly on current levels of factors (e.g., interest rates, earnings per share, etc.)

This leads to an HJB equation, but one that is even more difficult to solve than for the ordinary terminal wealth problem, except in rare, simple cases.

According to Markowitz (2004):

"*(Expected utility) is a reasonable assumption in theory, given the von Neumann and Morgenstern and the Leonard J. Savage axiomatic justifications of expected utility. But in practice, few if any investors know their utility functions; nor do the functions which financial engineers and financial economists find analytically convenient necessarily represent a particular investor's attitude towards risk and return.*"

A Risk Sensitive ICAPM

Suppose the measure of performance is, for some $\beta \in (0, \infty)$:

$$V(x, y) \equiv \liminf_{t \to \infty} (-2/\beta) t^{-1} \ln E[\exp^{-(\beta/2) \ln X_t} | X_0 = x, Y_0 = y]$$

Applying a Taylor series expansion of this about $\beta = 0$ one sees that

$$V(x, y) = \lim_{t \to \infty} t^{-t} E[\ln X_t] - \frac{\beta}{4} \lim_{t \to \infty} \mathrm{Var}(\ln X_t) + O(\beta^2)$$

$$= \text{geometric growth rate} - \frac{\beta}{4} (\text{asymptotic variance}) + O(\beta^2)$$

$$= \text{risk adjusted growth rate}$$

This leads to an HJB equation that is sometimes easier to solve (see "Risk Sensitive Intertemporal CAPM, with Application to Fixed Income Management" by T. Bielecki and SRP, *IEEE Trans. Auto. Control* 49, March 2004).

Example Risk Sensitive ICAPM

Suppose $(\mu_i(y)) = a + Ay$, $(\sigma_i(y)) = \Sigma$, and $C(y) = c + Cy$, where a, A, Σ, c, C, and Λ are all constant vectors and matrices.

We seek a solution $(\rho, \Phi(y))$ of the Hamilton-Jacobi-Bellman equation

$$\rho = \Phi_y'(y)Cy + c'\Phi_y(y) + \frac{1}{2}\Big[-\frac{\beta}{2}\Phi_y'(y)\Lambda\Lambda'\Phi_y'(y) + \mathrm{tr}\Lambda'\Phi_{yy}'(y)\Lambda\Big]$$

$$-\inf_{\{\pi:\mathbf{1}'\pi=1\}}\Big[\frac{1}{2}\big(\frac{\beta}{2}+1\big)\pi'\Sigma\Sigma'\pi - \pi'(a+Ay) + \frac{\beta}{2}\pi'\Sigma\Lambda'\Phi_y(y)\Big], \quad \forall y$$

Under certain technical assumptions this has a solution with ρ equal to the maximum risk adjusted growth rate and given by an explicit formula that does not depend upon x or y.

The optimal π is given by an explicit **affine** function of the factor vector y with coefficients that depend, in part, upon the solution of a **continuous algebraic Ricatti equation**.

A Statistics-Free Computational Approach

- Do not estimate the asset covariance matrix and the other parameters in the various SDE's
- Instead assume the asset weights are affine functions of the factor levels
- The variables are the affine function coefficients
- The objective is to maximize the risk adjusted growth rate when applied to historical data
- Equivalently, minimize the variance of the portfolio's historical returns minus a constant times the mean of historical returns
- Upper and lower bound constraints on the asset weights can be added
- This is a quadratic optimization problem
- Same approach can be used to compute optimal constant rebalance strategies (i.e., Merton strategies)

Example 2 (continued): Stan's Retirement Funds

- Interest rate is the factor, plus three risky assets and bank account.
- Regressions applied to 1995-2000 data showed the factor significantly affected only bond fund returns.
- So constant weights were assumed for stock and social choice funds, but affine weights for the bond fund (four coefficients).
- When using 1995-2000 data the bond weight was always zero, so...
- Considered revised problem with just two assets: bond fund and bank.
- Optimal risk sensitive strategies do somewhat better than optimal rebalance strategies when applied to the same data, but...
- Optimal risk sensitive strategies developed from 1995-2000 data were somewhat inferior when applied to 2001-2005 data.

Mathematical Justification of the Statistics-Free Approach

Work by and Related to "Nonparametric Kernel-Based Sequential Investment Strategies" by L. Györfi, G. Lugosi, and F. Udina

- Only statistical assumption: the vector sequence of asset returns is stationary and ergodic
- In this context, the maximum geometric growth rate is achieved by a **log optimum portfolio** strategy.
- Unfortunately, to compute such strategies one needs to know the probability law underlying the asset return process.
- Surprisingly, one can construct a strategy that does just as well without knowing the underlying distribution.
- "Just as well" means the constructed strategy is **universal**, that is, it has the same geometric growth rate as the log optimum portfolio strategy.

The Universal Portfolio Construction Approach

Consider a finite number of **experts**, or mutual fund managers.

Each expert is characterized by two numbers:

- A **window length**, e.g., four periods
- A measure of **similarity**, e.g., 2%

Each expert operates as follows:

- After the latest period look for all the similar windows in history
- Consider all strategies where π is vector of proportions for periods at ends of such windows, but equal proportions are used in other periods
- Choose π so as to maximize the current portfolio value
- Use this maximizing π during the coming period

For the universal portfolio, the proportions next period will be a weighted average of the proportions of the experts; the better an expert's performance, the bigger the weight.

Concluding Remarks

- Various mathematical optimization techniques seem useful for indexes, mutual funds, and the like, and....
- Perhaps derivatives like Markowitz options, which are tied to optimal investment decisions, will some day be marketed to ordinary people, but....
- Optimization techniques do not work very well for portfolios of individual stocks (hedge fund managers probably know more about this than do academics).
- To remedy this, perhaps better statistical models of stocks, suitable for long term investment, are needed.
- And/or we need to better develop non-statistical methods for portfolio optimization.