Financial Optimization
ISE 347/447

Lecture 12

Dr. Ted Ralphs

Reading for This Lecture

- C&T Chapter 8

Parameter Estimation in Portfolio Optimization

- To use Markovitz portfolio optimization in practice, one has to estimate the parameters μ_i and Q_{ij} $(i = 1, \ldots, n)$ from historical data.
- For recent previous investment periods $[t_{k-1}, t_k]$ $(k = 1, \ldots, T)$ having the same length as the current investment period, we can compute the returns
$$r_k^i = \frac{S_{t_k}^i - S_{t_{k-1}}^i}{S_{t_{k-1}}^i}.$$
- These numbers can be seen as independent samples of the distribution of R_i.

Method 1: Sample returns and covariances

- A simple approach is to estimate μ_i as the sample average return of asset i over the T previous investment periods,

$$\widehat{\mu}_i = \frac{1}{T}\sum_{k=1}^{T} r_k^i.$$

- The covariance Q_{ij} can be estimated similarly as

$$\widehat{Q}_{ij} = \frac{1}{T-1}\sum_{k=1}^{T}(r_k^i - \widehat{\mu}_i)(r_k^j - \widehat{\mu}_j).$$

Method 2: Capital Asset Pricing Model (CAPM)

- A far better method for estimating μ and Q from historical data is to use a market index S^m as a benchmark.
- Samples

$$r_k^m = \frac{S_{t_k}^m - S_{t_{k-1}}^m}{S_{t_{k-1}}^m}$$

of the random market return R^m are computed the same way as for other assets.

- Likewise, the risk-free returns r_k^f over the same periods can be obtained.

The CAPM Model

The aim is to find a model

$$r_k^i = r_k^f + \beta_i(r_k^m - r_k^f) + \varepsilon_k^i$$

that explains the excess return of the i^{th} asset over the risk-free return as a sum of two parts:

- $\beta_i(r_k^m - r_k^f)$ is fully explained by asset i's correlation with the market,
- ε_k^i is independent of the market and due to an unknown level of *idiosyncratic risk* σ_i^2 associated with asset i.
  - We interpret ε_k^i $(k = 1, \ldots, T)$ as i.i.d. samples of a Gaussian random variable $\mathcal{E}^i \sim \mathbb{N}(0, \sigma_i^2)$.
  - This part of the return is also thought of as being independent between different assets.

Fitting the Model

- We choose the model that minimizes the sample estimate of the idiosyncratic risk,

$$\sigma_i^2 = \min_{\beta_i} \frac{1}{T-1} \sum_{i=1}^{n} \left(r_k^i - r_k^f - \beta_i (r_k^m - r_k^f) \right)^2,$$

$$\beta_i = \arg\min_{\beta_i} \frac{1}{T-1} \sum_{i=1}^{n} \left(r_k^i - r_k^f - \beta_i (r_k^m - r_k^f) \right)^2.$$

- Taking derivatives with respect to β_i, it is easy to check that

$$\beta_i = \frac{\sum_k (r_k^i - r_k^f)(r_k^m - r_k^f)}{\sum_k (r_k^m - r_k^f)^2}.$$

- Under the CAPM model, one then stipulates

$$R_i = r_f + \beta_i (R_m - r_f) + \mathcal{E}^i.$$

The Estimators

- Therefore,

$$\mu_i = r_f + \beta_i(\mathbb{E}[R^m] - r_f),$$

$$\begin{aligned} Q_{ij} &= \mathrm{Cov}(r_f + \beta_i(R^m - r_f) + \mathcal{E}^i, r_f + \beta_j(R^m - r_f) + \mathcal{E}^j) \\ &= \beta_i\beta_j\sigma^2(R^m) + \delta_{ij}\sigma_i\sigma_j, \end{aligned}$$

  where δ_{ij} is the Kronecker delta.

- Replacing $\mathbb{E}[R^m]$ and $\sigma^2(R^m)$ by their sample estimators

$$\widehat{\mu}_m = T^{-1}\sum_k r_k^m,$$

$$\widehat{\sigma}_m^2 = \frac{1}{T-1}\sum_k \left(r_k^m - \widehat{\mu}_m\right)^2,$$

  we obtain the estimators

$$\widehat{\mu}_i = r_f + \beta_i\left(\widehat{\mu}_m - r_f\right),$$

$$\widehat{Q}_{ij} = \beta_i\beta_j\widehat{\sigma}_m^2 + \delta_{ij}\sigma_i\sigma_j.$$

The Multifactor CAPM Model

- The CAPM model can be generalized by replacing the benchmark index S^m with multiple indices $S^{m_1}, \dots, S^{m_q}$ for which samples

$$r_k^{m_j} = \frac{S_{t_k}^{m_j} - S_{t_{k-1}}^{m_j}}{S_{t_{k-1}}^{m_j}}$$

are available.

- The aim is now to find the best-fitting linear model

$$r_k^i = r_k^f + \sum_{j=1}^{q} \beta_i^j (r_k^{m_j} - r_k^f) + \varepsilon_k^i,$$

in full analogy to the above model.

Fitting the Multifactor Model

- We again choose the model that minimizes the sample estimate of the idiosyncratic risk,

$$\sigma_i^2 = \min_{(\beta_i^1,\ldots,\beta_i^q)} \frac{1}{T-1}\sum_{i=1}^{n}\Big(r_k^i - r_k^f - \sum_{j=1}^{q}\beta_i^j(r_k^{m_j} - r_k^f)\Big)^2,$$

$$(\beta_i^1,\ldots,\beta_i^q) = \arg\min_{(\beta_i^1,\ldots,\beta_i^q)} \frac{1}{T-1}\sum_{i=1}^{n}\Big(r_k^i - r_k^f - \sum_{j=1}^{q}\beta_i^j(r_k^{m_j} - r_k^f)\Big)^2.$$

- This is a strictly convex quadratic optimization problem when $T \geq q$, so the optimizer $(\beta_i^1,\ldots,\beta_i^q)$ can be computed by solving the linear system

$$\nabla_{(\beta_i^1,\ldots,\beta_i^q)}\sum_{i=1}^{n}\Big(r_k^i - r_k^f - \sum_{j=1}^{q}\beta_i^j(r_k^{m_j} - r_k^f)\Big)^2 = 0. \qquad (1)$$

The Estimators

- Let $B = (\beta_i^j)$ be the matrix of coefficients obtained when solving (1) for $(i = 1, \ldots, n)$.
- The multifactor CAPM model stipulates that

$$R = r_f e + B(P - r_f e) + \mathcal{E},$$

where $P = [P_1 \ \ldots \ P_q]^\top$ is the unknown random vector of returns

$$P_j = \frac{S_1^{m_j} - S_0^{m_j}}{S_0^{m_j}}$$

of the chosen indices over the investment period $[0, 1]$.

The Estimators (cont.)

- Therefore,

$$\mu = r_f e + B(\mathbb{E}[P] - r_f e),$$
$$Q = B\,\mathrm{Cov}(P) B^\top + \mathrm{Diag}(\sigma_1^2, \ldots, \sigma_n^2),$$

where $\mathrm{Cov}(P)$ is the variance-covariance matrix of P.

- Replacing $\mathbb{E}[P]$ and $\mathrm{Cov}(P)$ by their sample estimates yields the CAPM estimators for μ and Q.

The Black-Litterman Model

- This is really a method for estimating the vector μ of expected returns.
- The resulting estimates can be used in conjunction with any portfolio optimization model.
- As before, we assume an investment universe consists of n risky assets $S^1, \ldots, S^n$ and one risk-free asset S^0 with returns $R^1, \ldots, R^n$ (random) and r_f (deterministic) respectively over the investment period $[0, 1]$.
- We write $\tilde{R} = \begin{bmatrix} r_f \\ R \end{bmatrix}$, $\tilde{\mu} = \begin{bmatrix} r_f \\ \mu \end{bmatrix}$ and

$$\tilde{Q} = \mathrm{Cov}(\tilde{R}) = \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix},$$

where $Q = \mathrm{Cov}(R)$.

Key Assumptions

1. All investors use the investment strategy that maximizes

$$\text{(M)} \max_{\tilde{x} \in \mathbb{R}^{n+1}} \tilde{\pi}^\top \tilde{x} - \lambda \tilde{x}^\top \tilde{Q} \tilde{x}$$

$$\text{s.t.} \quad \sum_{i=0}^{n} \tilde{x}_i = 1,$$

   where $\tilde{\pi} = \left[\begin{smallmatrix} r_f \\ \pi \end{smallmatrix}\right]$ is an unknown vector of expected returns.

2. λ and Q are known (in reality this means they have already been estimated via some method of choice).

3. μ is itself modelled as a random vector $\mu = \pi + \nu$, where π is a return vector implied by the market and $\nu \sim \mathrm{N}(0, \tau Q)$ for some small τ

Key Assumptions (cont.)

The motivation for Assumption 3 is that if $R \sim \mathrm{N}(\mu, Q)$ and $R^1, \ldots, R^T$ are T independent samples of R, then the sample mean

$$\widehat{\mu} = \frac{1}{T} \sum_{i=1}^{T} R^i$$

has distribution $\mathrm{N}(\mu, T^{-1} Q)$.

Features of the Model

- The key distinguishing feature of the model is to allow investors to *specify their beliefs* about the performance of certain portfolios.
- This information is taken into account by updating the a priori estimate π of μ in a Bayes-like fashion to obtain an improved estimator.
- A key concept is the *return implied by the market*.
- Let $\tilde{w}$ be the relative weights of the assets in the market capitalization, i.e.,
$$\tilde{w}_i = \frac{z_i S^i}{\sum_{i=0}^{n} z_i S^i},$$
 where z_i is the number of shares of asset i that exist in the market and S^i the value of each share.

The Return Implied by the Market

- If Assumption 1 holds, $\tilde{w}$ must be the maximizer of (M).
- Otherwise the market would not be at equilibrium and prices would quickly adjust.
- Mathematically, this implies that $\tilde{w}$ has to satisfy the KKT conditions of (M),

$$\tilde{\pi} - 2\lambda\tilde{Q}\tilde{w} - \eta e^{[n+1]} = 0 \tag{2}$$

$$\sum_{i=0}^{n} \tilde{w}_i = 1. \tag{3}$$

- Here η is an unknown Lagrange multiplier, and $e^{[n+1]}$ is the $n+1$-dimensional vector of ones.

The Return Implied by the Market (cont.)

- Recalling that by Assumption 2, λ and $\tilde{Q}$ are known, we can solve for π.
- Since
$$\tilde{Q} = \mathrm{Cov}(\tilde{R}) = \begin{pmatrix} 0 & 0 \\ 0 & Q \end{pmatrix},$$
the first line of (2) reads $\eta = r_f$.
- Therefore,
$$\pi = r_f e^{[n]} + 2\lambda Q w.$$
- Having computed π, we now know the a priori distribution of μ,
$$\mu \sim \mathrm{N}(\pi, \tau Q).$$
- This is the *return implied by the market*.

Beliefs About the Market

- Market equilibrium is never quite reached because investors process new information at different speeds.
- Fund managers who process information faster than others therefore want the expected returns to reflect their insight or beliefs about where the market is headed next.
- For example, a manager who believes that the return of asset i is expected to outperform that of asset j by 2% would want μ to satisfy the constraint

$$\mu_i - \mu_j = 0.02.$$

Modeling Market Beliefs

- More generally, the fund manager may have several such beliefs concerning the expected returns of certain portfolios constructed from the risky assets $S^1, \ldots, S^n$, resulting in the constraints

$$A\mu = b,$$

where A is a $k \times n$-matrix and b a k-dimensional vector.

- A natural way of incorporating such beliefs into the parameter estimation problem is to use the conditional expectation

$$\widehat{\mu} = \mathbb{E}\left[\mu \mid A\mu = b\right]$$

as the estimator of the expected returns of assets $S^1, \ldots, S^n$.

Conditioning on Beliefs

- The conditional density of μ given $A\mu = b$ is

$$f(\mu \mid A\mu = b) = \frac{\exp\{-\frac{1}{2}(\mu - \pi)^\top (\tau Q)^{-1}(\mu - \pi)\}}{\int_{\{x : Ax = b\}} \exp\{-\frac{1}{2}(x - \pi)^\top (\tau Q)^{-1}(x - \pi)\} dx}$$

- Thus, the conditional distribution $\mathcal{D}(\mu \mid A\mu = b)$ is a multivariate Gaussian centered at $\widehat{\mu}$, which is where $f(\mu \mid A\mu = b)$ is maximized.

- Therefore, $\widehat{\mu}$ is the minimizer of the convex QP

$$\text{(P)} \quad \min_{\mu} \, (\mu - \pi)^\top (\tau Q)^{-1}(\mu - \pi) \quad \text{s.t.} \quad A\mu = b.$$

The Estimators

- Thus, $\widehat{\mu}$ is determined by the KKT conditions of problem (P):

$$(\tau Q)^{-1}(\widehat{\mu} - \pi) = A^\top \eta \tag{4}$$

$$A\widehat{\mu} = b, \tag{5}$$

  where η is a vector of Lagrange multipliers.

- From (4) we get

$$\widehat{\mu} = \pi + \tau Q A^\top \eta, \tag{6}$$

  and from (5), $b = A\widehat{\mu} = A\pi + \tau A Q A^\top \eta$.

- Hence,

$$\eta = (\tau A Q A^\top)^{-1}(b - A\pi)$$

  and

$$\widehat{\mu} = \pi + Q A^\top (A Q A^\top)^{-1}(b - A\pi). \tag{7}$$

Strength of Beliefs

- In reality, fund managers are never quite certain about their predictions.
- Therefore, we can also add randomness to the belief system $A\mu = b$ by expressing it as

$$A\mu = b + \mathcal{E}, \tag{8}$$

where $\mathcal{E} \sim \mathrm{N}(0, D)$ is a multivariate normal with zero mean and diagonal covariance matrix $D = \mathrm{Diag}(\sigma_1^2, \ldots, \sigma_n^2)$.

- In other words, the uncertainties pertaining to different constraints are considered to be independent.
- The larger the parameter σ_i^2, the less certain the investor is about the belief encoded in the i^{th} row of the system (8).

Incorporating Uncertain Beliefs

- Beliefs of the form $A\mu = b + \mathcal{E}$ cannot be conditioned upon, as they do not express an event.

- With some messy algebra, we can show that the generalized posterior distribution of μ is

$$\mathcal{D}(\mu \mid A\mu = b + \mathcal{E}) = \mathrm{N}\left(\widehat{\mu}, [(\tau Q)^{-1} + A^\top D^{-1} A]^{-1}\right),$$

where

$$\widehat{\mu} = \left[(\tau Q)^{-1} + A^\top D^{-1} A\right]^{-1} \left[(\tau Q)^{-1}\pi + A^\top D^{-1} b\right].$$

is the generalized posterior estimator of μ.

- Note that as uncertainty about the beliefs grows, $\widehat{\mu}$ tends to the a priori estimator π,

$$\lim_{D\to\infty} \widehat{\mu} = [(\tau Q)^{-1}]^{-1}(\tau Q)^{-1}\pi = \pi,$$

as one would expect.