IEOR 4004
Introduction to Portfolio Optimization

References: The classical reference is *Portfolio Selection: Efficient Diversification of Investments,* by Harry Markowitz. A more modern reference is: *Modern Portfolio Theory and Investment Analysis,* by Elton, Gruber, Brown and Goetzmann, ISBN 0471238546.

The general ob jective in static portfolio optimization is to make investments in assets (e.g. stocks) so that the overall portfolio provides good return, that is to say, it has a high growth rate per dollar invested while providing good protection against "risk". Typically, the total amount of money to be invested will be large, and, at the end of our selection process, the quantity of each asset that we actually invest in will also be large (or, rather, will not be tiny). In order to guide us in this process, all we have is a time series of data for each asset: its price history for, say, the past three years (for experimental purposes, we might instead simulate past data, e.g. using the geometric Brownian motion model). Thus, at first blush what we want is to pick a portfolio whose expected return is large. Experience says that we should also consider the standard deviation of the return, and in the end we will use an ob jective that trades-off expected return and standard deviation (the risk component).

Consider any given asset, i. Let its (closing) price at day t be $p_i^{(t)}$. If we own k_i units of this asset at the end of day t, then by the end of day $t+1$ our investment will have become $p_i^{(t+1)} k_i$. In other words, the return on asset i, on day t, (i.e. how much we gain, per dollar invested at day t) is:

$$\mu_i^{(t)} \;=\; \frac{p_i^{(t+1)} k_i - p_i^{(t)} k_i}{p_i^{(t)} k_i} \;=\; \frac{p_i^{(t+1)} - p_i^{(t)}}{p_i^{(t)}} \tag{1}$$

Suppose we were to observe the behavior of this asset (and others) over T days, say. Then we can compute its average return,

$$\bar{\mu}_i \;=\; \frac{1}{T}\left(\sum_{t=1}^{T} \mu_i^{(t)}\right). \tag{2}$$

Let's consider the portfolio implications of these quantities. Suppose we have a total budget of B dollars to invest on the assets, and that there are N assets in total. Our task is to choose, for each asset i, an amount w_i to invest on this asset. What can we say about the w_i ? Assuming (for simplicity) that we are considering long positions only,

$$w_i \geq 0, \;\text{ for } 1 \leq i \leq N. \tag{3}$$

Also,

$$\sum_{i=1}^{N} w_i \;=\; B. \tag{4}$$

The expected return of the portfolio is therefore

$$\frac{\sum_{i=1}^{N} \bar{\mu}_i \; w_i}{B}, \tag{5}$$

because the numerator is just the expected *change* in the value of the portfolio. If we use the notation

$$x_i = \frac{w_i}{B} \tag{6}$$

to indicate the *fraction* of money we invest in asset i, then the expected portfolio return can be rewritten as:

$$\sum_{i=1}^{N} \bar{\mu}_i x_i \tag{7}$$

Inequalities (3) and (4) can be rewritten as $x_i \geq 0$ and $\sum_i x_i = 1$, respectively. So, as a first attempt, we can view the static portfolio optimization problem as a **linear program**:

$$\begin{aligned} & \max \sum_{i=1}^{N} \bar{\mu}_i x_i \\ & \text{Subject to:} \\ & \sum_i x_i = 1, && (8) \\ & x \geq 0. && (9) \end{aligned}$$

It turns out that this is a very poor formulation of the problem. Among many weaknesses, it will invest all our budget on a single asset! We can try to control this by placing bounds on the assets for example, we might add a constraint stating that no asset can account for more than 5% of all the portfolio (e.g. $x_i \leq 0.05$ for all i). While such a constraint would help (and we do use such constraints in practice) the linear program has a deeper weakness: as we hold the portfolio over time, it will likely exhibit a huge variance, which is risk in other words. Solving the linear program gives us a portfolio with maximum expected return, but the portfolio will likely be very risky in that its value may well experience large losses.
What we need to do is to balance risk against return. This would lead us to look at an optimization problem of the form

$$\begin{aligned} & \min \lambda S(x) - \sum_{i=1}^{N} \bar{\mu}_i x_i \\ & \text{Subject to:} \\ & \sum_i x_i = 1, && (10) \\ & x \geq 0. && (11) \end{aligned}$$

where $S(x)$ is the standard deviation of the return of the portfolio (which is a function of the vector x) and $\lambda \geq 0$ is a parameter that we choose, that shows how we view the risk vs. return trade-off. Thus, when $\lambda = 0$, we are focusing just on return, while if $\lambda > 0$ is large, the emphasis is on minimizing $S(x)$, i.e. we are risk-averse. This is a sensible approach (though we will have to somehow choose λ) except that the optimization problem (10 - 12) is too hard! Instead, we will consider the objective function $\min \lambda S^2(x) - \sum_{i=1}^{N} \bar{\mu}_i x_i$, where $S^2(x)$ is the variance of the portfolio return. This problem will turn out to be tractable.

To understand these issues, let's compute $S^2(x)$. This equals:

$$\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}\mu_i^{(t)}x_i \;-\; \sum_{i=1}^{N}\bar{\mu}_i x_i\right)^2 \;=\; \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}x_i(\mu_i^{(t)}-\bar{\mu}_i)\right)^2. \tag{12}$$

If we expand this sum, we obtain

$$\sum_{i=1}^{N}x_i^2\left\{\frac{1}{T}\sum_{t=1}^{T}\left(\mu_i^{(t)}-\bar{\mu}_i\right)^2\right\} \;+\; 2\sum_{1\leq<j\leq N}x_ix_j\left\{\frac{1}{T}\sum_{t=1}^{T}\left(\mu_i^{(t)}-\bar{\mu}_i\right)\left(\mu_j^{(t)}-\bar{\mu}_j\right)\right\}. \tag{13}$$

The terms in (13) have standard statistical interpretations. For each asset i, denote

$$\sigma_i^2 \;=\; \frac{1}{T}\sum_{t=1}^{T}\left(\mu_i^{(t)}-\bar{\mu}_i\right)^2 \tag{14}$$

which is the variance of return of asset i, and for $i \neq j$ let

$$\sigma_{ij} \;=\; \frac{1}{T}\sum_{t=1}^{T}\left(\mu_i^{(t)}-\bar{\mu}_i\right)\left(\mu_j^{(t)}-\bar{\mu}_j\right) \tag{15}$$

which is the covariance of the returns of assets i and j (so $\sigma_{ij} = \sigma_{ji}$). In terms of these quantities, we obtain an optimization problem whose objective is:

$$\min \lambda\left(\sum_i \sigma_i^2 x_i^2 \;+\; 2\sum_{i<j}\sigma_{ij}x_ix_j\right) \;-\; \sum_i \bar{\mu}_i x_i \tag{16}$$

and with constraints given by equations (10) and (11). This is a (convex) quadratic programming problem, for which standard, efficient solution methods exist, as opposed to one where $S(x)$ appears directly in the ob jective (i.e., the *square root* of expression (16)).

0.1 Other constraints

Typical portfolio optimization problems will incorporate constraints other than (11) and (12). First of all, in general, we will have more significant bounds on the amounts that we can invest on any asset. For example, we might have a constraint stating:

$$0.001 \;\leq\; x_1 \;\leq\; 0.05, \tag{17}$$

which states that at most 5% of all the budget can be invested on asset 1, and that at least 0.1% of the budget must be invested on this asset. In general, for any asset i, we will have a "box" constraint of the form $d_i \leq x_i \leq u_i$. Here, the d and u values are parameters chosen by the portfolio manager. As an example of a more general constraint, suppose that assets 1 through 100 represent investments in a certain economic sector (e.g., high technology) and that we want to limit our exposure to this sector to at most 10% of our portfolio. In this case, we would state that:

$$\sum_{i=1}^{100}x_i \;\leq\; 0.1.$$

As a more complicated example, suppose assets 1 through 10 represents a set of companies involved in several business sectors, and that each dollar we invest in asset i translates into ai dollars of activity in, say, the energy sector. If we had a goal of investing at least 18% of our portfolio into the energy sector, we would express this goal with the constraint

$$\sum_{i=1}^{10} a_i x_i \ \geq \ 0.18.$$

We can expect that a complex portfolio optimization problem will have many such linear constraints. Furthermore, we may alter the ob jective, for example to try to encourage or discourage particular assets. Such a problem belongs to the general class of "convex quadratic programs":

$$\begin{aligned}
\min x'Qx \quad &- \quad c'x && (18)\\
\text{Subject to:} & && \\
A^{(1)}x \quad &\leq \quad b^{(1)} && (19)\\
A^{(2)}x \quad &\geq \quad b^{(2)} && (20)\\
A^{(3)}x \quad &= \quad b^{(3)} && (21)\\
d \leq \quad &x \quad \leq u. && (22)
\end{aligned}$$

Here, c is a vector with N entries (the "'" indicates transpose, so that $c'x = \sum_i c_i x_i$). Q is an $N \times N$ matrix (in our case, the matrix of covariances). Inequalities (22 - 24) are general linear constraints, in matrix form. Finally (25) indicates bounds on the individual variables.

0.2 How to solve quadratic programs

A problem of the form given above, where N is in the thousands and there are many con- straints (22 - 24), requires a sophisticated algorithm and a very good implementation in order to yield efficient (i.e., fast and robust) results. The book by Cornuéjols and Tütüncü contains some background. You can also consult *Introduction to Operations Research* by Hillier and Lieberman, or (for deeper material) *Linear and Nonlinear Programming*, by Luenberger.