PORTFOLIO OPTIMIZATION AND EXPECTED SHORTFALL MINIMIZATION FROM HISTORICAL DATA

We begin by describing the problem at hand which motivates our results. Suppose that we have \( n \) financial instruments at hand, each of whose price in one year is determined by some standard procedure \( \mathcal{P} \) practiced by the market, i.e. procedure \( \mathcal{P} \) is used to value each instrument today. We have an historical record of the prices of these instruments in the past. This set of data consists of a number \( d \) of fixed dates \( t_1, \ldots, t_d \) (in years) in the past, fixed dates being the same for all instruments, the value of each instrument at \( t_i + 1 \) calculated at time \( t_i \) using procedure \( \mathcal{P} \), and the value of each instrument historically attained at time \( t_i + 1 \), for all \( i \in [d] \).

We want to use these data to compare pricing via using procedure \( \mathcal{P} \) with actual real prices attained in the market, i.e. we can use the data to analyze how to make a profit today out of any discrepancies. Let \( p^\text{real}_{ij} \) be the actual price of instrument \( j \) at time \( t_i + 1 \) and let \( p^\mathcal{P}_{ij} \) be its price at \( t_i + 1 \) calculated at time \( t_i \) with \( \mathcal{P} \). Define

\[
    r_{ij} := \frac{p^\text{real}_{ij} - p^\mathcal{P}_{ij}}{p^\mathcal{P}_{ij}}, \quad \text{for all } i \in [d].
\]

The value of \( r_{ij} \) denotes the percent profit out of having bought a unit of instrument \( j \) at time \( t_i \) and then having sold it at time \( t_i + 1 \).

Investors today want to form a portfolio \( A \) from the \( n \) instruments and demand an expected (minimal) percent return \( r(A) = q \) in one year. Historical data may then be used to find one such portfolio, but the possible choices are usually infinite. One then prefers one portfolio \( A_1 \) over some other portfolio \( A_2 \) with \( r(A_1) = r(A_2) = q \) if we can argue that \( A_1 \) yields smaller risk than \( A_2 \). By smaller risk we probably understand something along the lines of: deviations from \( q \) are less likely to happen with portfolio \( A_1 \) than with portfolio \( A_2 \). The first problem is to define an adequate mathematical concept of risk. Several definitions are possible, but for our purposes we will use the idea of expected shortfall with cap 5 percent.

With risk now defined, we can formulate a mathematical problem out of our situation. As said, we want to find a portfolio \( A^* \) which yields a (minimal) return \( q \) and minimizes risk. Let \( w_j \in [0, 1] \) be the
percentage of investor’s money allocated to instrument \( j \in [n] \), which we call the weight of \( j \). The expected percent return of instrument \( j \) is given by

\[
r_j := \frac{\sum_{i=1}^{d} r_{ij}}{d}.
\]

Then, the expected percent return of a portfolio with weights \( w_1, \ldots, w_n \) is given by

\[
\sum_{j=1}^{n} r_j w_j.
\]

In our problem, the condition on the return is then given by

\[
\sum_{j=1}^{n} r_j w_j = q.
\]

On the other hand, the condition on the portfolio is stated as

\[
\sum_{j=1}^{n} w_j = 1 \text{ and } w_j \geq 0 \text{ for all } j \in [n].
\]

Clearly, we require \( \min\{r_1, r_2, \ldots, r_n\} \leq q \leq \max\{r_1, r_2, \ldots, r_n\} \) in order for all of this to make sense.

These conditions (equations 0.1 and 0.2) define a nonempty feasible region \( \mathcal{C} \) where we want to optimize. Let’s see what we can say about this set \( \mathcal{C} \). Equation 0.2 defines the standard \((n - 1)\)-dimensional hypersimplex in \( \mathbb{R}^n \), that is, the convex hull of the standard basis vectors of \( \mathbb{R}^n \):

\[
e_1 = (1, 0, 0, \ldots, 0),
\]
\[
e_2 = (0, 1, 0, 0, \ldots, 0),
\]

\[\vdots\]
\[
e_n = (0, 0, 0, \ldots, 0, 1).
\]

This set is usually denoted by \( \Delta_{n-1} \).

When we add the restriction of Equation 0.1, we obtain a linear cut of \( \Delta_{n-1} \). The data \( r_1, r_2, \ldots, r_n \) is obtained experimentally and so we may treat it as a set of \( n \) generic data. On the other hand, the value \( q \) is chosen by the investor in some interesting non-trivial way. Hence, we may assume that we have \( r_1 < r_2 < r_k < q < r_{k+1} < \cdots < r_n \) for some \( k \in [n] \) (or at least we may slightly vary the data to make it fit our model at no significant cost in the solution). The feasible region \( \mathcal{C} \)
of our problem is then an \((n - 2)\)-dimensional polytope in \(\mathbb{R}^n\) given as the convex hull of \(k(n - k)\) pairwise different vertices

\[ v_{1,k+1}, v_{1,k+2}, \ldots, v_{1,n}, \]
\[ v_{2,k+1}, v_{2,k+2}, \ldots, v_{2,n}, \]
\[ \vdots \]
\[ , v_{k,k+1}, v_{k,k+2}, \ldots, v_{k,n}, \]

and we can explicitly find these vertices at almost no computational cost:

For \(a \in [1, k]\) and \(b \in [k + 1, n]\) integers, we have

\[ v_{a,b} = \frac{r_b - q}{r_b - r_a} e_a + \frac{q - r_a}{r_b - r_a} e_b. \]

\[ = \left( 0, \ldots, 0, \frac{(r_b - q)/(r_b - r_a)}{\text{in position } a}, 0, \ldots, 0, \frac{(q - r_a)/(r_b - r_a)}{\text{in position } b}, 0, \ldots, 0 \right). \]

Denote this set of vertices by \(V\).

If the objective function \(f : C \to \mathbb{R}\) that we want to optimize (which accounts for risk) were linear, then we would be almost done with solving the linear optimization problem because we know the vertices of the feasible region. However, this is not the case in general.

On the good side, we will now argue that if we follow the definition of our problem, the objective function \(f\) (i.e. the negative of expected shortfall) that we want to optimize (maximize) will be a piecewise linear concave function so, in particular, it has a unique maximal value in \(C\).

Suppose that we fix some weights \(w = (w_1, \ldots, w_n) \in C\). From the set of \(d\) historical data, we obtain \(d\) historical percent returns of a portfolio with weights \((w_1, \ldots, w_n)\). From these data, we pick the worst 5 percent and calculate its average, and this is the value of \(f\) at \((w_1, \ldots, w_n)\). Our problem is to find those weights \(w^* = (w_1^*, \ldots, w_n^*) \in C\) at which \(f\) is maximized, so that the risk is minimized.

Fix some generic \(w \in C\). The average of the worst 5 percent historically on a portfolio with weights \(w\) comes from picking some dates in the past and then calculating the average percent return on those dates. By slightly varying the entries of \(w\) to obtain some \(w + \delta \in C\), so that \(|\delta|\) is very small, the same set of dates should again give the worst 5 percent for weights \(w + \delta\). Hence, we obtain that on a neighborhood \(U\) in \(C\) of \(w\), our objective function \(f\) is linear, so it coincides in \(U\) with some linear function \(h_w : C \to \mathbb{R}\). The existence of such neighborhood \(U\) is guaranteed by the fact that the number of historical data is finite and \(w\) is generic. At some non-generic points, however, we will change the choice of historical dates that give the worst 5 percent and this might change the choice of linear function \(g_1\) with which \(f\) coincides,
say, we change from $g_1$ to $g_2$. This change is due to $g_2$ yielding even worst values than $g_1$ on the average of the worst 5 percent. This reasoning shows that $f$ is actually a minimum of linear functions defined on $C$. Then, $f$ is a piecewise linear concave function on $C$ and so it has a unique maximum in our region.

So, now we have a very thorough understanding of the problem that we want to solve at almost no computational cost. We want to maximize a piecewise linear concave function $f$ on a region $C \subseteq \mathbb{R}^n$, and $C$ is an $(n-2)$-dimensional polytope all of whose vertices $V$ are known at almost no computational cost, i.e. $C$ is the set of all convex combinations of elements of $V$.

We can illustrate our results in the case where $n = 3$. We have simulated the percent return of three financial instruments on 100 dates, using the following parameters for the simulation:

- **Instrument 1:**
  - Expected percent return: 0.1.
  - Standard deviation: 0.12.

- **Instrument 2:**
  - Expected percent return: 0.05.
  - Standard deviation: 0.1.

- **Instrument 3:**
  - Expected percent return: 0.12.
  - Standard deviation: 0.13.

The specified value $q$ of expected return is 0.08. Using the results of the simulation, we have found the (necessarily) two vertices of the feasible region $C \subseteq \mathbb{R}^3$, they are:

$$v = (0.476891, 0.523109, 0) \quad \text{and} \quad u = (0, 0.553061, 0.446939).$$

All convex combinations $x = av + (1 - a)u$ of $v$ and $u$ produce all possible portfolios that yield an expected percent return $q = 0.08$, where the weight associated to instrument $i$ is given by the $i$-th entry of $x$.

Graphing Expected Shortfall ES against the coefficient $a$ in

$$av + (1 - a)u$$

we obtaining Figure 1. Recall that our objective function $f$ is always the negative of ES, so in our graph we have actually plotted $-f$. The minimum is attained approximately at $a = 0.524$, so that $x = av + (1 - a)u$ at this $a$ is:

$$x = (0.249891, 0.537366, 0.212743).$$
Figure 1. The function $-f$ (i.e. expected shortfall) for a simulated example with $n = 3$, which is then a piecewise linear convex function that we want to minimize.
This means that in an optimal portfolio, i.e. one with minimal risk, of percent return 8 percent, one needs to allocate 24.99 percent of the funds to Instrument 1, 53.74 percent to Instrument 2, and 21.27 percent to Instrument 3.