IEOR 4739
Further optimization tasks

A version of portfolio rebalancing. Suppose that on day t we own quantity q_j of asset j for $1 \leq j \leq n$. As before the current price (on day t) of asset j equals $p_j^{(t)}$. Thus the current fraction (percentage) of the portfolio that we own, in asset j, equals

$$\tilde{x}_j \doteq \frac{p_j^{(t)} q_j}{\sum_{i=1} p_i^{(t)} q_i}.$$

We want to buy/sell assets so as to convert these fractions into other fractions that are as close as possible to a previously computed portfolio given by fractions x_j^*. How do we do so?
Denote by d_j the number of shares that we buy/sell of asset j. So what we want is: .

$$p_j^{(t)}(q_j + d_j) = x_j^* \sum_{i=1}^{n} p_i^{(t)}(q_i + d_i) \quad 1 \leq j \leq n, \tag{1a}$$

$$q_j + d_j \geq 0, \quad 1 \leq j \leq n \tag{1b}$$

This basic set of equations describes the needed changes. But there may be many ways to attain (1) – hence we can talk about optimization.
a. Minimize value change needed to attain rebalancing.
Here we would solve the problem

$$\min \quad \sum_{i=1}^{n} p_i^{(t)}(d_i^+ + d_i^-) \tag{2a}$$

Subject to:

$$d_i = d_i^+ - d_i^-, \quad 1 \leq i \leq n, \tag{2b}$$

$$\text{Constraints (1).} \tag{2c}$$

b. Minimize (weighted)f number of shares needed to attain rebalancing.
Here instead we would solve the problem

$$\min \quad \sum_{i=1}^{n} W_i(d_i^+ + d_i^-) \tag{3a}$$

Subject to:

$$d_i = d_i^+ - d_i^-, \quad 1 \leq i \leq n, \tag{3b}$$

$$\text{Constraints (1).} \tag{3c}$$

In this case the W_i are given weights. For example if $W_i = 1$ for all i we are simply minimizing the number of traded shares. But in general we would have different weights – used to model difficulty in trading different assets.

The first two problems, in any case, may give rise to a large amount of trading. Rather than aiming for perfect rebalancing, we might instead ask for near-perfect rebalancing. As an example, suppose we are given (small) values $\varepsilon_i > 0$ for all i. Then we simply ask that

$$\left| \frac{p_j^{(t)} q_j}{\sum_{i=1} p_i^{(t)} q_i} - x_j^* \right| \leq \varepsilon_j$$

for all j. This allows some errors (relative to the x_j^*) in the final positions. Instead of (1) we will now have the inequalities

$$
\begin{aligned}
p_j^{(t)}(q_j + d_j) &\le (x_j^* + \varepsilon_j)\sum_{i=1}^{n} p_i^{(t)}(q_i + d_i) \quad 1 \le j \le n && \text{(4a)}\\
-p_j^{(t)}(q_j + d_j) &\le (-x_j^* + \varepsilon_j)\sum_{i=1}^{n} p_i^{(t)}(q_i + d_i) \quad 1 \le j \le n && \text{(4b)}\\
q_j + d_j &\ge 0, \quad 1 \le j \le n && \text{(4c)}
\end{aligned}
$$

Using this criterion we obtain problems similar to to those considered in (a) and (b) above.

First version of robust portfolio optimization. In previous lectures we considered the classical mean-variance portfolio optimization problem, which in the long-only variant takes the form

$$
\begin{aligned}
&\min \; \lambda\, x^T Q x \;-\; \sum_{i=1}^{N} \mu_j x_j \\
&\text{Subject to:} \\
&\sum_j x_j = 1 && \text{(5a)}\\
&Ax \ge b && \text{(5b)}\\
&x \ge 0. && \text{(5c)}
\end{aligned}
$$

Here $\lambda \ge 0$ is the risk aversion parameter, Q is the covariance of returns, μ is the vector of average asset returns, and $Ax \ge b$ are structural constraints. Note that $Ax \ge b$ can be used to model many requirements, including bounds on the x_j.

What we are interested in here is the behavior of the portfolio when the μ_j are noisy estimates of the "true" expected asset returns, that is to say, they include errors. Thus, we run the risk of constructing a portfolio that will behave *very* differently when the real returns are realized – even if the real returns are slightly different from the values μ_j.

We can study this problem by imagining that future average returns will be set by an imaginary adversary. This is best considered as a two-step process:

(1) First, a decision-maker (you) picks the asset weights x_j.

(2) Second, the adversary picks an expected returns vector $\hat{\mu}$. This vector is chosen from some set $\mathscr{M} \subseteq \mathbb{R}^n$. The goal of the adversary is to *minimize* the return earned by x, that is to say, the adversary solves the optimization problem

$$
r(x) \quad \doteq \quad \min \sum_{i=1}^{n} \hat{\mu}_j x_j \;:\; \hat{\mu} \in \mathscr{M}. \tag{6}
$$

Here, $r(x)$ is the *worst-case* expected return earned by the weights x, given the set $\mathscr{M}$. Using these ideas, we can now formulate a "robust portfolio optimization" problem, which is the variant of the mean-variance problem where you protect yourself against asset average return mis-estimations. This updated problem is as follows

$$
\begin{aligned}
\min \lambda x^T Q x \quad &- \quad \rho && \text{(7a)}\\
\text{Subject to:} & && \\
\sum_j x_j \quad &= \quad 1 && \text{(7b)}\\
Ax \quad &\geq \quad b && \text{(7c)}\\
x \quad &\geq \quad 0 && \text{(7d)}\\
\rho \quad &\leq \quad \sum_j \hat{\mu}_j x_j \qquad \forall \hat{\mu} \in \mathscr{M} && \text{(7e)}
\end{aligned}
$$

In this formulation, ρ is an added variable that represents the value $r(x)$ as per equation (6). Constraint (7e)) makes sure that indeed, at optimality we will have $\rho = r(x)$ (can you see why?)

There remains the task of solving the optimization problem (7). In general this is a challenging task because this is a nonstandard optimization problem – it is an optimization problem with another optimization problem as a constraint. This is called a bilevel problem.

However, there are cases where problem (7) can be reduced to a standard problem. An interesting case arises when the set $\mathscr{M}$ is as follows. First we have nonnegative constants $\gamma_1, \gamma_2, \ldots, \gamma_n$, and an additional constant Γ. Then $\hat{\mu} \in \mathscr{M}$ whenever

(i) First, $|\hat{\mu}_j - \mu_j| \leq \gamma_j$ for $1 \leq j \leq n$.

(ii) Second, $\sum_{j=1}^n |\hat{\mu}_j - \mu_j| \leq \Gamma$.

Rule (i) states that none of the return estimates is "very" wrong – the error in asset j is at most γ_j. Rule (ii) states that the sum of all errors is also not very large.

Example. Suppose $n = 3000$, $\gamma_j = 0.001$ for all j and that $\Gamma = 0.1$. Then, for example, you could have 100 assets each exhibiting an error $|\hat{\mu}_j - \mu_j| = 0.001$ – but all other assets would have to have no errors at all, i.e. they would have to have $|\hat{\mu}_j - \mu_j| = 0$. Or, you could have 1000 assets each attaining $|\hat{\mu}_j - \mu_j| = 0.0001$. Or, you could have every asset attaining an error $|\hat{\mu}_j - \mu_j| = 0.1/3000$. Or many other combinations that satisfy (i) and (ii).

Next we show to solve problem (7) as a standard optimization problem. To this end, let us write the problem solved by the adversary, *given a vector of asset weights* x, as follows

$$
\begin{aligned}
r(x) \ = \ \min \quad & \sum_{j=1}^n x_j \hat{\mu}_j && \text{(8a)}\\
\text{Subject to:} & && \\
\delta_j - \hat{\mu}_j \quad &\geq \quad -\mu_j \quad 1 \leq j \leq n && \text{(8b)}\\
\delta_j + \hat{\mu}_j \quad &\geq \quad \mu_j \quad 1 \leq j \leq n && \text{(8c)}\\
0 \leq \delta_j \quad &\leq \quad \gamma_j \quad 1 \leq j \leq n && \text{(8d)}\\
\sum_{j=1}^n \delta_j \quad &\leq \quad \Gamma. && \text{(8e)}
\end{aligned}
$$

In this formulations the only variables are the $\hat{\mu}_j$ and the δ_j – everything else is data. Note that (8b) and (8c) imply $\delta_j \geq |\hat{\mu}_j - \mu_j|$. In fact, because of (8e), *without loss of generality* at optimality we will have $|\hat{\mu}_j - \mu_j| = \delta_j$ for each j.

What we have attained so far is to represent the worst case adversarial return, $\rho(x)$, as the value of a linear program. But how do we use this fact to convert (7) into a standard problem? Answer: we use LP-duality. Denote the dual variable associated with (8b), (8c), (8d) and (8e) by p_j, q_j, $-s_j$ and $-t$, respectively.

Then, using LP duality, we can write LP (8) in the equivalent form

$$
\begin{aligned}
r(x) \ = \ \max \quad & \sum_{j=1}^{n} \{-\mu_j p_j + \mu_j q_j - \gamma_j s_j + \Gamma t\} && \text{(9a)} \\
\text{Subject to:} \quad & && \\
& p_j + q_j - s_j - t \ \leq \ 0, \quad 1 \leq j \leq n && \text{(9b)} \\
& -p_j + q_j \ \leq \ x_j, \quad 1 \leq j \leq n && \text{(9c)} \\
& p, q, s \geq 0, \quad t \geq 0. && \text{(9d)}
\end{aligned}
$$

So now we can rewrite the robust problem (7) in the following form:

$$
\begin{aligned}
\min \lambda x^T Q x \quad & - \quad \rho && \text{(10a)} \\
\text{Subject to:} \quad & && \\
\sum_j x_j \quad & = \quad 1 && \text{(10b)} \\
Ax \quad & \geq \quad b && \text{(10c)} \\
x \quad & \geq \quad 0 && \text{(10d)} \\
\rho \quad & \leq \quad \sum_{j=1}^{n} \{-\mu_j p_j + \mu_j q_j - \gamma_j s_j + \Gamma t\} && \text{(10e)} \\
& p_j + q_j - s_j - t \ \leq \ 0, \quad 1 \leq j \leq n && \text{(10f)} \\
& -p_j + q_j \ \leq \ x_j, \quad 1 \leq j \leq n && \text{(10g)} \\
& p, q, s \geq 0, \quad t \geq 0. && \text{(10h)}
\end{aligned}
$$

In this problem the variables are the x_j, ρ, as well as the p_j, q_j, s_j and t. Do you see why this formulation works?

A more general case. We now take up problem (7) in the more general case of the set $\mathscr{M}$ where for appropriate matrix C and vector d,

$$
\mathscr{M} \quad = \quad \{\hat{\mu} : C\hat{\mu} \geq d\}. \tag{11}
$$

The specific case considered above, being linear, falls under this category. We will see that in this general case there is a formulation which, like (10), is "tractable". First let us write the adversarial problem (for a given portfolio vector x) in dual form

$$
\begin{aligned}
r(x) \quad & \doteq \quad \min \sum_{i=1}^{n} \hat{\mu}_j x_j \ : \ \hat{\mu} \in \mathscr{M}, && \text{(12a)} \\
& = \quad \min \sum_{i=1}^{n} \hat{\mu}_j x_j && \text{(12b)} \\
& \qquad \text{s.t.} \quad C\hat{\mu} \ \geq \ d && \\
& = \quad \max d^T \alpha && \text{(12c)} \\
& \qquad \text{s.t.} \quad C^T \alpha \ = \ x, && \\
& \qquad \qquad \alpha \geq 0. &&
\end{aligned}
$$

In this formulations, the maximization in (12c) yields $r(x)$ because of linear programming duality. We can now write the robust portfolio optimization problem

$$
\begin{aligned}
\min\ \lambda\, x^T Q x \quad &- \quad \rho && (13a)\\
\text{Subject to:} & \\
\sum_j x_j \quad &= \quad 1 \\
Ax \quad &\geq \quad b \\
x \quad &\geq \quad 0 \\
\rho \quad &\leq \quad \sum_j \hat{\mu}_j x_j \qquad \forall\, \hat{\mu} \in \mathscr{M} && (13b)
\end{aligned}
$$

$=$

$$
\begin{aligned}
\min\ \lambda\, x^T Q x \quad &- \quad \rho && (13c)\\
\text{Subject to:} & \\
\sum_j x_j \quad &= \quad 1 \\
Ax \quad &\geq \quad b \\
x \quad &\geq \quad 0 \\
\rho \quad &\leq \quad d^T \alpha && (13d)\\
C^T \alpha - x \quad &= \quad 0, \\
\alpha \quad &\geq \quad 0.
\end{aligned}
$$

We explain the validity of this statement next – but note that the formulation beginning in (13c) is a standard (convex) quadratic programming problem, with variables x, ρ and α, and so it can be solved with any standard solver. Now to see why the formulation is valid, suppose that (x^*, ρ^*, α^*) is *feasible* for this problem. Then weak LP duality implies that $\rho^* \leq r(x^*)$ (because α^* is feasible for the dual, i.e. α^* is feasible for the formulation beginning with (12c) when $x = x^*$). If in addition (x^*, ρ^*, α^*) is *optimal* we would in fact set $\rho^* = r(x^*)$ since and not $\rho^* < r(x^*)$ since ρ only appears in constraint (13d) and the objective of the problem in (13c) or (13a) wants to make ρ as large as can be, subject to all the other constraints. So, indeed, the formulation beginning in (13c) is a correct representation of the problem in (12)

Exercise. Show that the formulation in (10) amounts to an application of these ideas.

Wed.Mar.30.205348.2016@rabbitchaser