ORF 307: Lecture 18

Parametric Simplex Method
and
The Efficient Frontier

Robert J. Vanderbei

April 23, 2019

Slides last edited on January 25, 2019

Recall Ingredients for Portfolio Optimization

Raw Data:

$$R_j(t) \quad = \quad \text{return on asset } j \text{ in time period } t$$

$\Longrightarrow$

Derived Data:

$$r_j = \frac{1}{T}\sum_{t=1}^{T} R_j(t)$$

$$D_{tj} = R_j(t) - r_j.$$

Decision Variables:

$$x_j \quad = \quad \text{fraction of portfolio to invest in asset } j$$

Decision Criteria:

$$\text{reward}(x) = \sum_j r_j x_j$$

$$\text{risk}(x) = \frac{1}{T}\sum_{t=1}^{T}\left|\sum_j D_{tj}x_j\right|$$

Optimization Problem

Set a value for *risk affinity* parameter μ (risk affinity is the reciprocal of risk aversion)

and maximize a combination of *reward minus risk*:

$$\text{maximize} \quad \mu \sum_j r_j x_j - \frac{1}{T} \sum_{t=1}^{T} \left| \sum_j D_{tj} x_j \right|$$

$$\text{subject to} \quad \sum_j x_j = 1$$

$$x_j \geq 0 \qquad \text{for all investments } j$$

Because of absolute values not a linear programming problem.

As usual, easy to convert...

A Linear Programming Formulation

$$\begin{aligned}
\text{maximize} \quad & \mu \sum_j r_j x_j - \frac{1}{T} \sum_{t=1}^{T} y_t \\
\text{subject to} \quad & -y_t \leq \sum_j D_{tj} x_j \leq y_t \qquad \text{for all times } t \\
& \sum_j x_j = 1 \\
& x_j \geq 0 \qquad \text{for all investments } j \\
& y_t \geq 0 \qquad \text{for all times } t
\end{aligned}$$

Note: The y_t's are the absolute values of the deviations from the average reward. To be clear: they are *not the dual variables.*

Adding Slack Variables w_t^+ and w_t^-

$$
\begin{array}{llll}
\text{maximize} & \mu \sum_j r_j x_j - \frac{1}{T} \sum_{t=1}^{T} y_t & & \\
\text{subject to} & -y_t - \sum_j D_{tj} x_j + w_t^- = 0 & & \text{for all times } t \\
 & -y_t + \sum_j D_{tj} x_j + w_t^+ = 0 & & \text{for all times } t \\
 & \sum_j x_j = 1 & & \\
 & x_j \geq 0 & & \text{for all investments } j \\
 & y_t,\ w_t^-,\ w_t^+ \geq 0 & & \text{for all times } t
\end{array}
$$

A dictionary will have $2T+1$ equations and $3T+n$ variables, $2T+1$ of which will be basic and rest will be nonbasic (here, n denotes the number of assets).

The Solution for Large μ

Varying the risk bound $0 \leq \mu < \infty$ produces the *efficient frontier*.

Large values of μ favor reward maximization whereas small values favor minimizing risk.

Beyond some finite (but perhaps large) value for μ, the optimal solution will be a portfolio consisting of just one asset—the asset j^* with the largest average return:

$$r_{j^*} \geq r_j \qquad \text{for all } j.$$

For this case, it's easy to identify *basic* (nonzero) vs. *nonbasic* (i.e. zero) variables:

- Variable x_{j^*} is basic whereas the remaining x_j's are nonbasic.
- All of the y_t's are basic.
- If $D_{tj^*} > 0$, then w_t^- is basic and w_t^+ is nonbasic. Otherwise, it is switched.

The algebra is tedious, but we can now write down a starting dictionary...

The Optimal Dictionary for Large μ

Let

$$T^+ = \{t : D_{tj^*} > 0\}, \quad T^- = \{t : D_{tj^*} < 0\}, \quad \text{and} \quad \epsilon_t = \begin{cases} 1, & \text{for } t \in T^+ \\ -1, & \text{for } t \in T^- \end{cases}$$

Here's the optimal dictionary (for μ large):

$$\begin{array}{lcllllll}
\zeta & = & \frac{1}{T}\sum_{t=1}^{T} \epsilon_t D_{tj^*} & -\frac{1}{T}\sum_{j \neq j^*}\sum_{t=1}^{T} \epsilon_t (D_{tj} - D_{tj^*}) x_j & -\frac{1}{T}\sum_{t \in T^-} w_t^- & -\frac{1}{T}\sum_{t \in T^+} w_t^+ & \\
 & & +\mu r_{j^*} & +\mu \sum_{j \neq j^*} (r_j - r_{j^*}) x_j & & & \\
\hline
y_t & = & -D_{tj^*} & -\sum_{j \neq j^*} (D_{tj} - D_{tj^*}) x_j & +w_t^- & & t \in T^- \\
w_t^- & = & 2D_{tj^*} & +2\sum_{j \neq j^*} (D_{tj} - D_{tj^*}) x_j & & +w_t^+ & t \in T^+ \\
y_t & = & D_{tj^*} & +\sum_{j \neq j^*} (D_{tj} - D_{tj^*}) x_j & & +w_t^+ & t \in T^+ \\
w_t^+ & = & -2D_{tj^*} & -2\sum_{j \neq j^*} (D_{tj} - D_{tj^*}) x_j & +w_t^- & & t \in T^- \\
x_{j^*} & = & 1 & -\sum_{j \neq j^*} x_j & & &
\end{array}$$

An Example

Collected data for 719 stocks (and bonds, etc.) from January 1, 1990, to March 18, 2002.

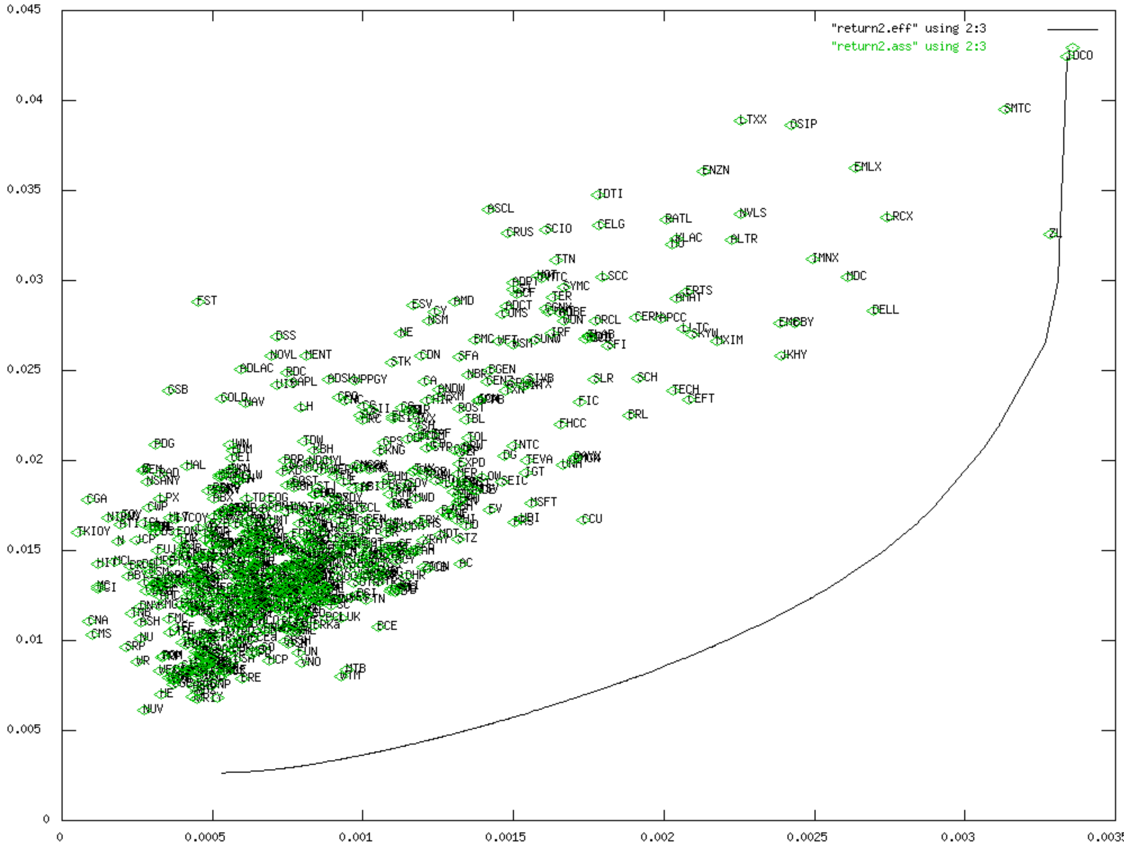
Hence,

$$n = 719$$

and

$$T = 3080.$$

Efficient Frontier



Click here for an expanded browser view.

Computing the Efficient Frontier

Using a reasonably efficient code for the parametric self-dual simplex method (simpo), it took *22,000* pivots and *1.5 hours* to solve for *one point* on the efficient frontier.

Customizing this same code to solve parametrically for every point on the efficient frontier, it took *20,500* pivots and *57 minutes* to compute *every point* on the frontier.

The efficient frontier consists of *1308* distinct portfolios. Click here for a complete list (*warning: the file is 2.5 MBytes*).

A Different Application — Sparse Regression

Lasso Regression

The problem is to solve a sparsity-encouraging "regularized" regression problem:

$$\text{minimize} \quad \|Ax - b\|_2^2 + \lambda \|x\|_1$$

My reaction:

Why not replace *least squares* (LS) with *least absolute deviations* (LAD)?

LAD is to LS as median is to mean. Median is a more robust statistic (i.e., insensitive to outliers).

The LAD version can be recast as a *linear programming* (LP) problem.

If the solution is expected to be sparse, then the *simplex method* can be expected to solve the problem very quickly.

No one knows the "correct" value of the parameter λ. The *parametric simplex method* can solve the problem for *all values of* λ from $\lambda = \infty$ to a small value of λ in the same (fast) time it takes the standard simplex method to solve the problem for one choice of λ.