Project Report

Jesse Lu

August 14, 2013

1 Main result: Quasi pareto-optimal portfolios

n_{signals}	1	2	4	8	16	32	64	128
	4.519	4.548	4.591	4.819	5.501	6.271	6.848	7.182
	4.075	3.574	3.500	3.922	4.876	5.767	6.406	6.768
	1.187	2.290	2.497	3.138	4.339	5.263	5.909	6.275
	1.241	1.824	2.127	2.869	4.086	5.020	5.654	6.011
	1.252	1.749	2.051	2.794	4.018	4.952	5.583	5.939

Table 1: Annualized returns of portfolios in figure 1.

n_{signals}	1	2	4	8	16	32	64	128
	0.348	0.375	0.425	0.513	0.641	0.739	0.801	0.837
	0.533	0.661	0.848	1.149	1.500	1.864	2.054	2.161
	0.421	0.679	1.083	1.691	2.664	3.968	4.650	4.965
	0.416	0.655	0.999	1.759	3.466	6.304	8.300	8.933
	0.413	0.654	0.980	1.756	3.656	6.972	9.504	10.200

Table 2: Information ratios of portfolios in figure 1.

n_{signals}	1	2	4	8	16	32	64	128
	0.163	0.159	0.151	0.142	0.148	0.163	0.175	0.183
	0.131	0.123	0.116	0.114	0.130	0.153	0.170	0.179
	0.061	0.084	0.082	0.095	0.122	0.149	0.168	0.178
	0.058	0.073	0.071	0.090	0.120	0.148	0.168	0.177
	0.058	0.071	0.069	0.089	0.119	0.148	0.167	0.177

Table 3: Turnover of portfolios in figure 1.

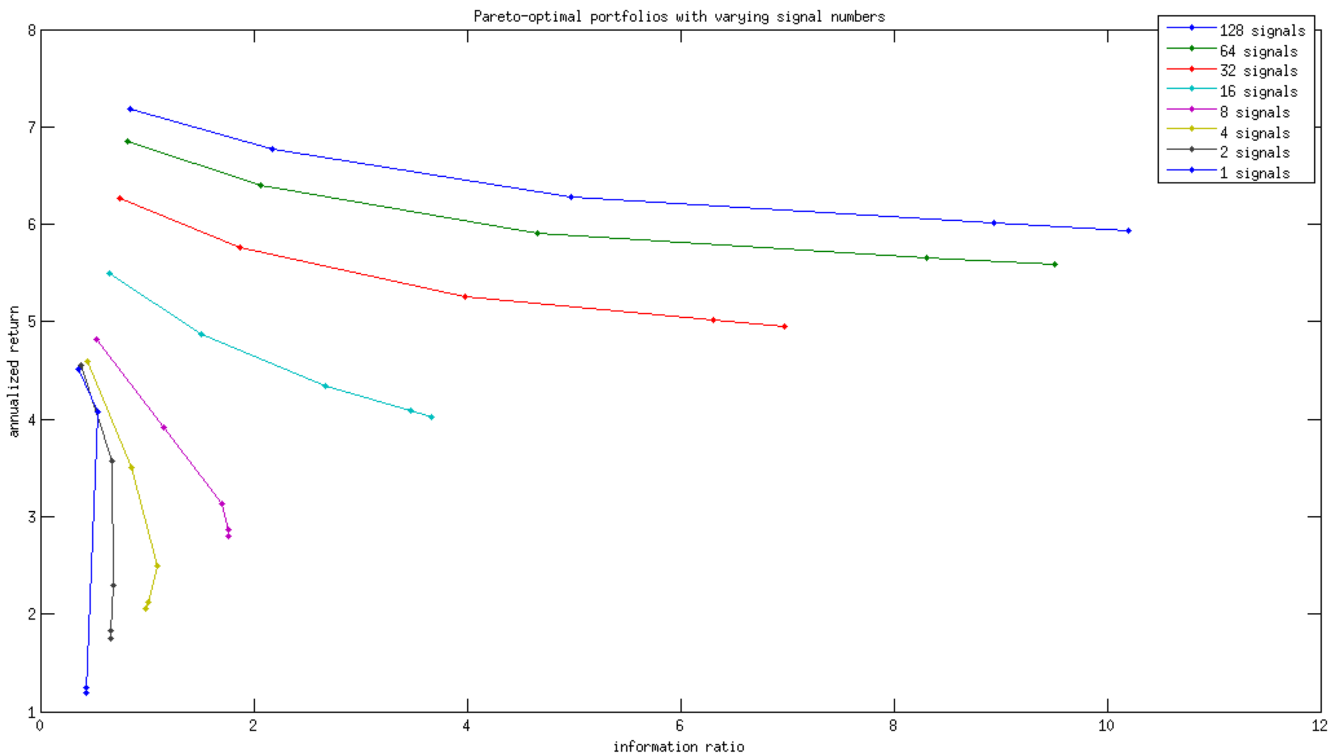

Figure 1: The main result of the project: quasi pareto-optimal portfolios with varying numbers of signals, where even small signal numbers can produce large returns with large information ratios. Here, each point represents a portfolio, while lines connect portfolios generated by models of the same number of signals. The variety of portfolios for each signal number group is due to differentiated weighting between maximizing either the annualized return or the information ratio.

2 Finding the near-optimal forward-biased portfolio

2.1 Problem statement

The portfolio optimization problem can be simplified by substituting expressions for the relevant terms in the following way (implemented in `simulate_X.m`):

$$f_{ir}(x) = \frac{c^T x}{\text{std}(Ax)} \tag{1a}$$

$$f_{ret}(x) = a\frac{c^T x}{\|x\|_1} \tag{1b}$$

$$f_{tvr}(x) = \text{mean}\frac{C|Bx|}{C|x|} \tag{1c}$$

$$f_t(x) = \frac{\|Bx\|_1}{\|x\|_1} \tag{1d}$$

where for an m stocks by n days portfolio,

- $x \in \mathbf{R}^{mn}$ is a vector representing the portfolio,
- c, a, A, B and C are appropriate vector, scalar, or matrix quantities (see `sim_matrices.m`), and
- $f_{ir}, f_{ret}, f_{tvr}, f_t \in \mathbf{R}^{mn} \to \mathbf{R}$ where f_t is an approximation of the turnover function, f_{tvr}.

The problem we want to solve is then

$$\begin{aligned} \text{minimize} \quad & -f_{ret}(x) - \hat{\mu} f_{ir}(x) && \text{(2a)} \\ \text{subject to} \quad & f_{tvr}(x) \leq 0.2 && \text{(2b)} \end{aligned}$$

That is to say, maximize the return and information ratio (with relative proportion affected by coefficient $\hat{\mu}$) while keeping the turnover below 0.2.

Note that for brevity, I have not included the expressions for c, a, A, B and C; however, one can deduce that such quantities must exist. For example, notice that $c^T x$ represents the total profit/loss; we know this expression must be valid for some c since the expression for the total profit/loss is linear. Lastly, notice that the portfolio is represented as a vector (x), rather than a matrix, in this section for notational convenience.

2.2 Problem simplification

We now simplify (2) in order to make it computationally tractable.

We first note that (2) is homogenous, meaning that it's solution depends only on the relative values of x. Therefore we can split the terms of $f_{tvr}(x)$ in the following way without loss of generality.

$$\begin{aligned} \text{minimize} \quad & -f_{ret}(x) - \hat{\mu} f_{ir}(x) && \text{(3a)} \\ \text{subject to} \quad & \|Bx\|_1 \leq 0.2 && \text{(3b)} \\ & \|x\|_1 = 1 && \text{(3c)} \end{aligned}$$

Turning our attention to the objective function we note that

$$f_{ret}(x) + \hat{\mu} f_{ir}(x) = \frac{c^T x}{\text{std}(Ax)} + \hat{\mu} a c^T x \tag{4}$$

since $\|x\|_1 = 1$. This objective can be thought of as weighting two terms by $\hat{\mu}$; and we can alternatively weight the following two terms

$$\frac{c^T x}{\text{std}(Ax)} + \hat{\mu} a c^T x \to (1/\mu) c^T x - \text{std}(Ax) \tag{5}$$

while still obtaining the full range of pareto-optimal points; although the connection between μ and $\hat{\mu}$ is not immediately clear.

This now gives us

$$\begin{aligned} \text{minimize} \quad & (-1/\mu) c^T x + \text{std}(Ax) && \text{(6a)} \\ \text{subject to} \quad & \|Bx\|_1 \leq 0.2 && \text{(6b)} \\ & \|x\|_1 = 1 && \text{(6c)} \end{aligned}$$

which is still completely equivalent to (2).

We now implement the following generalization

$$\|x\|_1 = 1 \quad \to \quad \|x\|_1 \leq 1 \tag{7}$$

in order to make the problem convex[1]. Our reasoning behind this approximation is that although $\|x\|_1$ is no longer strictly confined to the value of 1, it most likely will end up with a quantity of 1 anyways because the $c^T x$ term in the objective function will always want to "push" $\|x\|_1$ to as large a value as possible.

While this approximation does inject error into our formulation, the benefits of convexity (i.e. the ability to efficiently compute the global minimum) far outweigh the costs. Actually, the cost of this approximation can be made negligible by twiddling the value of 0.2 in the $\|Bx\|_1 \leq 0.2$ constraint (using a simple algorithm like bisection for instance) in order to force $f_{tvr}(x) \leq 0.2$.

We have now simplified (2) into the following convex form

$$\begin{aligned} \text{minimize} \quad & (-1/\mu) c^T x + \text{std}(Ax) && \text{(8a)} \\ \text{subject to} \quad & \|Bx\|_1 \leq 0.2 && \text{(8b)} \\ & \|x\|_1 \leq 1 && \text{(8c)} \end{aligned}$$

[1] see Boyd and Vandenberghe, "Convex Optimization"