Optimization Assignment 4

Ze Yang (zey@andrew.cmu.edu)

September 24, 2018

Problem 1. **(a)** Selecting stocks with highest volatilities yield a selection of

- Non-Zero Indices: (4, 6, 13, 14, 18)
- Tickers: (DGX, STZ, NTAP, LH, XTO)
- Weights: (0.2219, 0.2182, 0.2443, 0.1283, 0.1873)
- Minimum tracking error: 698.193

(b) Forward selection yields:

- Non-Zero Indices: (1, 4, 7, 8, 17)
- Tickers: (NE, DGX, TIF, SVU, MKC)
- Weights: (0.1999, 0.1747, 0.1401, 0.2638, 0.2215)
- Minimum tracking error: 77.237

(c) Tree searching with forward selection pruning yields:

- Non-Zero Indices: (1, 4, 7, 8, 16)
- Tickers: (NE, DGX, TIF, SVU, FDO)
- Weights: (0.1968, 0.1896, 0.1368, 0.2854, 0.1914)
- Minimum tracking error: 72.792

Problem 2. **(a)** It suffices to show $x_i(\boldsymbol{Vx})_i$ is a constant (not a function of i). Let $\boldsymbol{\theta} = (1/\sigma_1 \ 1/\sigma_2 \ ... \ 1/\sigma_n)^\top$; then $\boldsymbol{x}$ can be rewritten as $\boldsymbol{x} = c\boldsymbol{\theta}$, where $c = 1/\sum_1^n (1/\sigma_k)$ is also a scalar constant. Hence

$$\boldsymbol{Vx} = c\left(\rho\boldsymbol{\sigma}\boldsymbol{\sigma}^\top + (1-\rho)\mathrm{Diag}\{\boldsymbol{\sigma}\}^2\right)\boldsymbol{\theta} = (1-\rho+n\rho)\, c\boldsymbol{\sigma} \tag{1}$$

$$x_i(\boldsymbol{Vx})_i = \frac{c}{\sigma_i}\cdot(1-\rho+n\rho)\, c\sigma_i = (1-\rho+n\rho)\, c^2 \quad \forall\ i = 1, ..., n \tag{2}$$

Therefore, $RC_i(\boldsymbol{x}) = RC_j(\boldsymbol{x}) = (1-\rho+n\rho)\, c^2/\sqrt{\boldsymbol{x}^\top\boldsymbol{Vx}}$.

(b) Recall from homework 3, by Woodbury formula:

$$\boldsymbol{V}^{-1} = \frac{1}{1-\rho}\mathrm{Diag}(\boldsymbol{\theta})^2 - \frac{\rho}{(1-\rho)(1-\rho+n\rho)}\boldsymbol{\theta}\boldsymbol{\theta}^\top \tag{3}$$

It's easy to verify:

$$\boldsymbol{V}^{-1}\boldsymbol{\sigma} = \left(\frac{1}{1-\rho} - \frac{n\rho}{(1-\rho)(1-\rho+n\rho)}\right)\boldsymbol{\theta} = \frac{1}{1-\rho+n\rho}\boldsymbol{\theta} \tag{4}$$

Also note that the maximum diversification problem has exactly the same mathematical form as the maximum Sharpe ration problem, with $\boldsymbol{\sigma}$ in place of $\boldsymbol{\mu}$. By the solution of homework 3, the solution to it is:

$$\boldsymbol{x}^* = \frac{1}{\boldsymbol{1}^\top\boldsymbol{V}^{-1}\boldsymbol{\sigma}}\boldsymbol{V}^{-1}\boldsymbol{\sigma} = \frac{\boldsymbol{\theta}}{\boldsymbol{1}^\top\boldsymbol{\theta}} = \frac{1}{\sum_{k=1}^n \frac{1}{\sigma_k}}\boldsymbol{\theta} \tag{5}$$

Which is exactly: $\boldsymbol{x}_i^* = \frac{1}{\sum_{k=1}^n \frac{1}{\sigma_k}}\frac{1}{\sigma_i}$

(3) By definition of risk parity:

$$RC_i(\boldsymbol{x}) = \frac{x_i(\boldsymbol{V}\boldsymbol{x})_i}{\sqrt{\boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{x}}} = \frac{\sqrt{\boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{x}}}{n} \tag{6}$$

Therefore, x_i satisfies:

$$x_i = \frac{\boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{x}}{n(\boldsymbol{V}\boldsymbol{x})_i} = \frac{1}{n\beta_i} \tag{7}$$

Since $\boldsymbol{x}$ is fully invested: $\boldsymbol{1}^\top \boldsymbol{x} = \sum_{k=1}^n \frac{1}{n\beta_k} = 1 \Rightarrow \sum_{k=1}^n \frac{1}{\beta_k} = n$. Therefore:

$$x_i = \frac{1}{n\beta_i} = \frac{1}{\sum_{k=1}^n \frac{1}{\beta_k}} \frac{1}{\beta_i} \tag{8}$$

Problem 3. (a) Denote the Lagrangian relaxation problem (LR); let the feasible region for (SC) and (LR) be $D_{\mathrm{SC}}, D_{\mathrm{LR}}$ respectively. We have $D_{\mathrm{SC}} \subseteq D_{\mathrm{LR}}$.
Define $f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u}) := \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} x_{ij} + \sum_{i=1}^n u_i \left(1 - \sum_{j=1}^n x_{ij}\right)$; clearly, $f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u})$ is affine in $\boldsymbol{u}$ for any $(\boldsymbol{x}, \boldsymbol{y}) \in D_{\mathrm{LR}}$. The Lagrange (partial) dual function can be written as, in terms of f:

$$L(\boldsymbol{u}) = \max_{\boldsymbol{x},\boldsymbol{y} \in D_{\mathrm{LR}}} f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u}) \tag{9}$$

Consider for $\lambda \in [0, 1]$, and arbitrary $(\boldsymbol{x}, \boldsymbol{y}) \in D_{\mathrm{LR}}$:

$$\begin{aligned} f_{\boldsymbol{x},\boldsymbol{y}}(\lambda \boldsymbol{u}_1 + (1-\lambda)\boldsymbol{u}_2) &\le \lambda f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u}_1) + (1-\lambda) f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u}_2) \qquad \text{(convexity of } f_{\boldsymbol{x},\boldsymbol{y}}) \\ &\le \max_{\boldsymbol{x}',\boldsymbol{y}' \in D_{\mathrm{LR}}} \left[\lambda f_{\boldsymbol{x}',\boldsymbol{y}'}(\boldsymbol{u}_1) + (1-\lambda) f_{\boldsymbol{x}',\boldsymbol{y}'}(\boldsymbol{u}_2)\right] \\ &\le \max_{\boldsymbol{x}',\boldsymbol{y}' \in D_{\mathrm{LR}}} \lambda f_{\boldsymbol{x}',\boldsymbol{y}'}(\boldsymbol{u}_1) + \max_{\boldsymbol{x}',\boldsymbol{y}' \in D_{\mathrm{LR}}} (1-\lambda) f_{\boldsymbol{x}',\boldsymbol{y}'}(\boldsymbol{u}_2) \\ &= \lambda L(\boldsymbol{u}_1) + (1-\lambda)\lambda L(\boldsymbol{u}_2) \end{aligned} \tag{10}$$

This holds for arbitrary $\boldsymbol{x}, \boldsymbol{y} \in D_{\mathrm{LR}}$, so we are free to take the maximum of left hand side over D_{LR}:

$$L(\lambda \boldsymbol{u}_1 + (1-\lambda)\boldsymbol{u}_2) \le \lambda L(\boldsymbol{u}_1) + (1-\lambda)\lambda L(\boldsymbol{u}_2) \tag{11}$$

Which is the desired result.

(b) Let $g(\boldsymbol{x}, \boldsymbol{y})$ be the objective function for (SC). Suppose $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) \in D_{\mathrm{SC}} \subseteq D_{\mathrm{LR}}$ is an arbitrary feasible point for (SC). We have: $f_{\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{y}}}(\boldsymbol{u}) - g(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) = \sum_{i=1}^n u_i \left(1 - \sum_{j=1}^n \widetilde{x}_{ij}\right) = 0$ for any $\boldsymbol{u}$. Therefore:

$$L(\boldsymbol{u}) = \max_{\boldsymbol{x},\boldsymbol{y} \in D_{\mathrm{LR}}} f_{\boldsymbol{x},\boldsymbol{y}}(\boldsymbol{u}) \ge f_{\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{y}}}(\boldsymbol{u}) = g(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) \tag{12}$$

Since this holds for arbitrary feasible point $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) \in D_{\mathrm{SC}}$, we are free to take the maximum of the right hand side:

$$L(\boldsymbol{u}) \ge \max_{\widetilde{\boldsymbol{x}},\widetilde{\boldsymbol{y}} \in D_{\mathrm{SC}}} g(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{y}}) = Z \tag{13}$$

(c) The only difference between D_{SC} and D_{LR} is the constraint $\sum_{j=1}^n x_{ij} = 1$. So clearly $\bar{\boldsymbol{y}}$ satisfies D_{SC}'s contraints. It suffices to check $\bar{\boldsymbol{x}}$. By definition, for each $i = 1, ..., n$, $\bar{x}_{ij} = 1$ for only the j with the highest ρ_{ij} (we can add some infinitesimal perturbation to ρ if there are ties). Thus $\sum_{j=1}^n \bar{x}_{ij} = 1$ for $i = 1, ..., n$. So $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \in D_{\mathrm{SC}}$. The final conclusion is obvious: If $L(\boldsymbol{u}) = g(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}}) \ge Z \ge g(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$, it's clear that the optimal value $Z = g(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$. Thus $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ is optimal.