Optimization Assignment 3

Ze Yang (zey@andrew.cmu.edu)

September 18, 2018

Problem 1.

(a) Let $\boldsymbol{u} = \sqrt{\rho}\boldsymbol{\sigma}$, $\boldsymbol{A} = (1-\rho)\text{diag}(\boldsymbol{\sigma})^2$. $\boldsymbol{u}^\top \boldsymbol{A}^{-1} = \frac{\sqrt{\rho}}{1-\rho}(1/\sigma_1 \, ... \, 1/\sigma_n) = \frac{\sqrt{\rho}}{1-\rho}\boldsymbol{\theta}^\top$
Then $\boldsymbol{V} = \boldsymbol{A} + \boldsymbol{u}\boldsymbol{u}^\top$. We have $1 + \boldsymbol{u}^\top \boldsymbol{A}^{-1}\boldsymbol{u} = 1 + n\rho/(1-\rho) \neq 0$ while $\rho \in (0,1)$. Therefore $\boldsymbol{V}$ is invertible. By the Woodbury formula:

$$\begin{aligned} \boldsymbol{V}^{-1} = (\boldsymbol{A} + \boldsymbol{u}\boldsymbol{u}^\top)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{u}^\top \boldsymbol{A}^{-1}}{1 + \boldsymbol{u}^\top \boldsymbol{A}^{-1}\boldsymbol{u}} &= \frac{1}{1-\rho}\text{diag}(1/\boldsymbol{\sigma})^2 - \frac{(\sqrt{\rho}/(1-\rho))^2 \boldsymbol{\theta}\boldsymbol{\theta}^\top}{1 + n\rho/(1-\rho)} \\ &= \underbrace{\frac{1}{1-\rho}}_{a}\text{diag}(\boldsymbol{\theta})^2 - \underbrace{\frac{\rho}{(1-\rho)(1-\rho+n\rho)}}_{b}\boldsymbol{\theta}\boldsymbol{\theta}^\top \end{aligned} \tag{1}$$

(b) The minimum-rike fully invested portfolio is given by the optimal solution $\boldsymbol{x}^* = \frac{1}{\boldsymbol{1}^\top \boldsymbol{V}^{-1}\boldsymbol{1}}\boldsymbol{V}^{-1}\boldsymbol{1} = \frac{1}{\boldsymbol{1}^\top \boldsymbol{y}}\boldsymbol{y} = \frac{1}{\sum_i y_i}\boldsymbol{y}$, if we let $\boldsymbol{y} := \boldsymbol{V}^{-1}\boldsymbol{1}$. It suffices to calculate $\boldsymbol{y}$. Denote $\boldsymbol{\theta} \circ \boldsymbol{\theta} := (1/\sigma_1^2 \, ... \, 1/\sigma_n^2)^\top$

$$\begin{aligned} \boldsymbol{y} &= (a \cdot \text{diag}(\boldsymbol{\theta})^2 - b \cdot \boldsymbol{\theta}\boldsymbol{\theta}^\top)\boldsymbol{1} = a \cdot (\boldsymbol{\theta} \circ \boldsymbol{\theta}) - b\sum_{j=1}^n \frac{1}{\sigma_j}\boldsymbol{\theta} \\ y_i &= \frac{a}{\sigma_i^2} - b\sum_{j=1}^n \frac{1}{\sigma_i \sigma_j} = \frac{a}{\sigma_i^2}\left(1 - \frac{b}{a}\sum_{j=1}^n \frac{\sigma_i}{\sigma_j}\right) \end{aligned} \tag{2}$$

Hence $c = a/\sigma_i^2 = \frac{1}{(1-\rho)\sigma_i^2}$, $d = b/a = \frac{\rho}{1-\rho+n\rho}$.

Problem 2.

(a) Let $\boldsymbol{z} := k\boldsymbol{x}$ such that $\boldsymbol{\mu}^\top \boldsymbol{z} = 1$. The maximum Sharpe problem is equivalent to:

$$\begin{aligned} \min_{\boldsymbol{z},k} \quad & \boldsymbol{z}^\top \boldsymbol{V}\boldsymbol{z} \\ s.t. \quad & \boldsymbol{\mu}^\top \boldsymbol{z} = 1 \\ & \boldsymbol{1}^\top \boldsymbol{z} = k \\ & -k \leq 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} &\mathcal{L}(\boldsymbol{z}, k; \boldsymbol{\lambda}) = \boldsymbol{z}^\top \boldsymbol{V}\boldsymbol{z} + \lambda_1(\boldsymbol{\mu}^\top \boldsymbol{z} - 1) + \lambda_2(\boldsymbol{1}^\top \boldsymbol{z} - k) + \lambda_3(-k) \\ &\begin{cases} \nabla_{\boldsymbol{z}}\mathcal{L} = \boldsymbol{V}\boldsymbol{z} + \lambda_1\boldsymbol{\mu} + \lambda_2\boldsymbol{1} = 0 \\ \nabla_k \mathcal{L} = -\lambda_2 - \lambda_3 = 0 \\ \lambda_3 k = 0; \quad \lambda_3, k \geq 0 \\ \boldsymbol{\mu}^\top \boldsymbol{z} = 1 \\ \boldsymbol{1}^\top \boldsymbol{z} = k \end{cases} \end{aligned} \tag{3}$$

Since there's at least one entry in $\boldsymbol{\mu}$ being strictly positive, $\boldsymbol{z}$ can't be all-zeros $\Rightarrow k \neq 0$. Hence the complementary slackness condition $\Rightarrow \lambda_3 = 0$. The gradient of k equation $\Rightarrow \lambda_2 = -\lambda_3 = 0$. The first equation $\Rightarrow \boldsymbol{V}\boldsymbol{z} + \lambda_1\boldsymbol{\mu} = 0 \Rightarrow \boldsymbol{z}^* = \lambda_1 \boldsymbol{V^{-1}\mu}$. The last equation $\Rightarrow \lambda_1 = k/\boldsymbol{1^\top V^{-1}\mu}$. Hence $\boldsymbol{z}^* = \frac{k}{\boldsymbol{1^\top V^{-1}\mu}}\boldsymbol{V^{-1}\mu}$. Consequently $\boldsymbol{x}^* = \boldsymbol{z}^*/k = \frac{1}{\boldsymbol{1^\top V^{-1}\mu}}\boldsymbol{V^{-1}\mu}$.

(b) Substitute $\boldsymbol{\mu}$ for $\delta \boldsymbol{V}\boldsymbol{x}_B$, $\boldsymbol{x}_B$ is also fully invested:

$$\boldsymbol{x}^* = \frac{1}{\boldsymbol{1^\top V^{-1}}\delta\boldsymbol{V}\boldsymbol{x}_B}\boldsymbol{V^{-1}}\delta\boldsymbol{V}\boldsymbol{x}_B = \frac{\boldsymbol{x}_B}{\boldsymbol{1}^\top \boldsymbol{x}_B} = \boldsymbol{x}_B \tag{4}$$

(c)

$$\begin{aligned} \min_{\boldsymbol{\mu}} \quad & (\boldsymbol{\pi}-\boldsymbol{\mu})^\top \boldsymbol{Q}^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}) \\ s.t. \quad & \boldsymbol{P}\boldsymbol{\mu}=\boldsymbol{q} \end{aligned} \quad \Rightarrow \quad \begin{aligned} & \mathcal{L}(\boldsymbol{\mu};\lambda) = (\boldsymbol{\pi}-\boldsymbol{\mu})^\top \boldsymbol{Q}^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}) + \lambda^\top(\boldsymbol{P}\boldsymbol{\mu}-\boldsymbol{q}) \\ & \begin{cases} \nabla_{\boldsymbol{\mu}}\mathcal{L} = -\boldsymbol{Q}^{-1}(\boldsymbol{\pi}-\boldsymbol{\mu}) + \boldsymbol{P}^\top\boldsymbol{\lambda} = \boldsymbol{0} \\ \boldsymbol{P}\boldsymbol{\mu}-\boldsymbol{q}=\boldsymbol{0} \end{cases} \end{aligned} \tag{5}$$

Condition (1) $\Rightarrow \widehat{\boldsymbol{\mu}} = \boldsymbol{\pi} - \boldsymbol{Q}\boldsymbol{P}^\top\boldsymbol{\lambda}^*$. Premultiply by $\boldsymbol{P} \Rightarrow \boldsymbol{P}\boldsymbol{\pi} - \boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}^\top\boldsymbol{\lambda}^* = \boldsymbol{q} \Rightarrow \boldsymbol{\lambda}^* = (\boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}^\top)^{-1}(\boldsymbol{P}\boldsymbol{\pi} - \boldsymbol{q})$. Therefore $\widehat{\boldsymbol{\mu}} = \boldsymbol{\pi} + \boldsymbol{Q}\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{Q}\boldsymbol{P}^\top)^{-1}(-\boldsymbol{P}\boldsymbol{\pi}+\boldsymbol{q})$.

(d) Plug in $\boldsymbol{\mu} = \widehat{\boldsymbol{\mu}}$, $\boldsymbol{Q} = \boldsymbol{V}$, $\boldsymbol{\pi} = \delta\boldsymbol{V}\boldsymbol{x}_B$:

$$\begin{aligned} \boldsymbol{x}^{**} &= \frac{1}{\boldsymbol{1}^\top\boldsymbol{V^{-1}}[\boldsymbol{\pi}+\boldsymbol{V}\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\pi})]}\boldsymbol{V^{-1}}\left[\boldsymbol{\pi}+\boldsymbol{V}\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}-\boldsymbol{P}\boldsymbol{\pi})\right] \\ &= \frac{\boldsymbol{V^{-1}}\left[\delta\boldsymbol{V}\boldsymbol{x}_B+\boldsymbol{V}\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}-\delta\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B)\right]}{\boldsymbol{1}^\top\boldsymbol{V^{-1}}[\delta\boldsymbol{V}\boldsymbol{x}_B+\boldsymbol{V}\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}-\delta\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B)]} \\ &= \frac{\boldsymbol{x}_B+\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}/\delta-\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B)}{\boldsymbol{1}^\top\boldsymbol{x}_B+\boldsymbol{1}^\top\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}/\delta-\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B)} \\ &= \lambda\boldsymbol{x}_B+\boldsymbol{P}^\top\boldsymbol{v} \end{aligned} \tag{6}$$

where

$$\begin{aligned} \lambda &= \frac{1}{\boldsymbol{1}^\top\boldsymbol{x}_B+\boldsymbol{1}^\top\boldsymbol{P}^\top(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}/\delta-\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B)} \\ \boldsymbol{v} &= \lambda(\boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^\top)^{-1}(\boldsymbol{q}/\delta-\boldsymbol{P}\boldsymbol{V}\boldsymbol{x}_B) \end{aligned} \tag{7}$$

Problem 3.

(a)

1. The curve of sorted eigenvalues of $\widehat{\boldsymbol{V}}$ is *steeper* than that of $\boldsymbol{V}$; that is, $\hat{\lambda}$'s overestimates the greater λ's of actual $\boldsymbol{V}$, while underestimates the smaller ones. See (Figure 1) below.
2. Actual variance > true optimal variace > estimated variance.
3. This pattern persists. See (Figure 2) below.

(b)

1. The curve of sorted eigenvalues of the shrinkage estimate $\bar{\boldsymbol{V}}$ is still steeper than that of $\boldsymbol{V}$, but not as much as that of $\hat{\boldsymbol{V}}$. See (Figure 1) below.
2. Actual variance > true optimal variace > estimated variance. However, the spread between actural variance - estimated variance is now narrower, compared with the naive sample estimate. In general, we have:

$$\hat{\boldsymbol{x}}^\top\boldsymbol{V}\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^\top\widehat{\boldsymbol{V}}\hat{\boldsymbol{x}} > \bar{\boldsymbol{x}}^\top\boldsymbol{V}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}^\top\bar{\boldsymbol{V}}\bar{\boldsymbol{x}} > 0 \tag{8}$$

   Or in another word:

$$\hat{\boldsymbol{x}}^\top\boldsymbol{V}\hat{\boldsymbol{x}} > \bar{\boldsymbol{x}}^\top\boldsymbol{V}\bar{\boldsymbol{x}} > \boldsymbol{x}^{*\top}\boldsymbol{V}\boldsymbol{x}^* > \bar{\boldsymbol{x}}^\top\bar{\boldsymbol{V}}\bar{\boldsymbol{x}} > \hat{\boldsymbol{x}}^\top\widehat{\boldsymbol{V}}\hat{\boldsymbol{x}} \tag{9}$$

3. This pattern persists. See (Figure 2) below.

Figure 1: Sorted Eigenvalues of True, Sample Estimate, and Shrinkage Estimate of Covariance Matrix

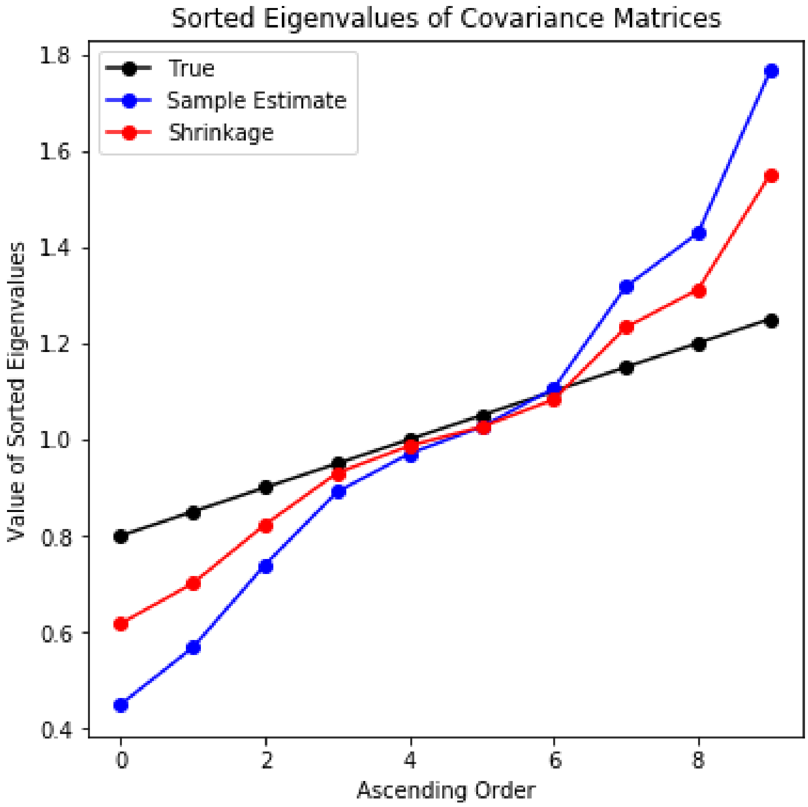

Figure 2: Estimated and Realized Variance of Minimum-Risk Portfolio