Exercises

Portfolio Optimization: Theory and Application
Chapter 13 – Index Tracking Portfolios

Daniel P. Palomar (2025). *Portfolio Optimization: Theory and Application.*
Cambridge University Press.

portfoliooptimizationbook.com

Exercise 13.1: Indices and ETFs

Download price data corresponding to some financial indices (e.g., the S&P 500, Dow Jones Industrial Average, Nasdaq) and some ETFs that track each of these indices (e.g., SPY for the S&P 500 index). Plot each index along with the corresponding ETFs in a linear and a logarithmic scale. Assess the tracking capabilities.

Exercise 13.2: Active vs. passive investments

Download price data corresponding to some mutual funds and compare with appropriate financial indices. Plot the price time series and compute some performance measure, such as the Sharpe ratio, to compare their performance. Do these results support the efficient-market hypothesis, promoted by Fama, or the inefficient and irrational markets, promoted by Shiller?

Exercise 13.3: Sparse regression via ℓ_1-norm

Generate an underdetermined system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{5\times 10}$. Then, solve the following sparse underdetermined system of linear equations via brute force (i.e., trying all possible 2^{10} patterns for the variable $\boldsymbol{x}$):

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{x}\|_0 \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

Finally, solve the following linear program and compare the solution with the previous one:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{x}\|_1 \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

Exercise 13.4: Sparse least squares

Generate an overdetermined system of linear equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with $\boldsymbol{A} \in \mathbb{R}^{10\times 5}$. Consider the resolution of the sparse regression problem

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \\ \text{subject to} & \|\boldsymbol{x}\|_0 \leq k \end{array}$$

via the following list of methods and plot the trade-off curve of regression error vs. sparsity level for each method:

a. Brute force (i.e., trying all possible 2^5 patterns for the variable $\boldsymbol{x}$).

b. ℓ_1-norm approximation:

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_1.$$

c. Concave approximation using a general-purpose nonlinear solver:

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda\sum_{i=1}^{N}\log\left(1 + \frac{|x_i|}{\varepsilon}\right).$$

d. Concave approximation again, but using the iterative reweighted ℓ_1-norm method.

Exercise 13.5: Cap-weighted indices

The portfolio of a cap-weighted index is defined in terms of the market capitalization. Denoting by $\boldsymbol{p}_t$ the prices of the N assets at time t and by $\boldsymbol{n}$ the number of outstanding shares of the N assets. The capital portfolio of the assets is defined to be proportional to the market capitalization $\boldsymbol{n} \odot \boldsymbol{p}_t$, which leads to the normalized portfolio

$$\boldsymbol{b}_t = \frac{\boldsymbol{n} \odot \boldsymbol{p}_t}{\boldsymbol{n}^\mathsf{T}\boldsymbol{p}_t}.$$

Show that this normalized portfolio can also be expressed as

$$\boldsymbol{b}_t = \frac{\boldsymbol{b}_{t-1} \odot (\mathbf{1} + \boldsymbol{r}_t)}{\boldsymbol{b}_{t-1}^\mathsf{T}(\mathbf{1} + \boldsymbol{r}_t)},$$

where the returns are defined as

$$\boldsymbol{r}_t = \frac{\boldsymbol{p}_t - \boldsymbol{p}_{t-1}}{\boldsymbol{p}_{t-1}} = \frac{\boldsymbol{p}_t}{\boldsymbol{p}_{t-1}} - \mathbf{1}.$$

Exercise 13.6: Tracking error measures

Download price data corresponding to some financial index (e.g., the S&P 500, Dow Jones Industrial Average, Nasdaq) and some ETFs that track the index (e.g., SPY for the S&P 500 index). Compute different error tracking measures, namely the ℓ_2-norm tracking error, the downside risk, the ℓ_1-norm tracking error, and the Huberized tracking error. Finally, plot a histogram of the tracking errors as a more complete picture of the tracking performance (note that the previous error measures are summarizations of the histogram).

Exercise 13.7: Two-stage index tracking methods

Download price data corresponding to some financial index, such as the S&P 500, and the corresponding constituent N assets for some period of time. Then, construct the benchmark return vector $\boldsymbol{r}^{\mathrm{b}}$ and the assets' return matrix $\boldsymbol{X}$, and formulate the sparse index tracking problem

$$\begin{array}{ll}\underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2\\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\; \boldsymbol{w} \geq \mathbf{0},\\ & \|\boldsymbol{w}\|_0 \leq K.\end{array}$$

a. Solve the problem via a naive two-stage approach: simply select the K active assets with some heuristic and then renormalize so that $\mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$.

b. Solve the problem via a two-stage approach with refitting of weights: select the K active assets as before and then solve the convex regression problem with the selected assets.

Plot the trade-off curve of regression error vs. sparsity level K for each method.

Exercise 13.8: Sparse index tracking methods via concave sparsity approximation

Download price data corresponding to some financial index, such as the S&P 500, and the corresponding constituent N assets for some period of time. Then, construct the benchmark return vector $\boldsymbol{r}^{\mathrm{b}}$ and the assets' return matrix $\boldsymbol{X}$, and formulate the sparse index tracking problem

$$\begin{array}{ll}\underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda\|\boldsymbol{w}\|_0\\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\; \boldsymbol{w} \geq \mathbf{0}\end{array}$$

for different values of the hyper-parameter λ.

a. Approximate the sparsity regularizer with the concave log-function:

$$\begin{array}{ll}\underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda\sum_{i=1}^{N}\log\left(1 + \frac{|w_i|}{\varepsilon}\right)\\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\; \boldsymbol{w} \geq \mathbf{0}.\end{array}$$

Then solve the problem with a general-purpose nonconvex solver.

b. Apply the majorization–minimization approach to get the iterative reweighted ℓ_1-norm method that solves sequentially, $k = 0, 1, 2, \dots$, the following:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda \sum_{i=1}^{N} \alpha_i^k |w_i| \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\ \boldsymbol{w} \geq \mathbf{0}, \end{array}$$

where

$$\alpha_i^k = \frac{1}{\varepsilon + \left|w_i^k\right|}.$$

Plot the trade-off curve of regression error vs. sparsity level for each method (by varying the hyper-parameter λ).

Exercise 13.9: Sparse index tracking for downside risk

Download price data corresponding to some financial index, such as the S&P 500, and the corresponding constituent N assets for some period of time. Then, construct the benchmark return vector $\boldsymbol{r}^{\mathrm{b}}$ and the assets' return matrix $\boldsymbol{X}$, and formulate the sparse index tracking problem

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\left(\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right)^+\right\|_2^2 + \lambda\|\boldsymbol{w}\|_0 \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\ \boldsymbol{w} \geq \mathbf{0} \end{array}$$

for different values of the hyper-parameter λ.

a. Approximate the sparsity regularizer with the concave log-function and solve the problem with a general-purpose nonconvex solver.

b. Apply the majorization–minimization approach to get the iterative reweighted ℓ_1-norm method that solves sequentially the following convex problem:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\left(\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right)^+\right\|_2^2 + \lambda \sum_{i=1}^{N} \alpha_i^k |w_i| \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\ \boldsymbol{w} \geq \mathbf{0}, \end{array}$$

where

$$\alpha_i^k = \frac{1}{\varepsilon + \left|w_i^k\right|}.$$

c. Apply the majorization–minimization approach fully to get the iterative reweighted ℓ_1-norm method that solves sequentially the following convex problem:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\left(\tilde{\boldsymbol{r}}^{\mathrm{b}}\right)^k - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda \sum_{i=1}^{N} \alpha_i^k |w_i| \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1,\ \boldsymbol{w} \geq \mathbf{0}, \end{array}$$

where now

$$\left(\tilde{\boldsymbol{r}}^{\mathrm{b}}\right)^k = \boldsymbol{r}^{\mathrm{b}} + \left(\boldsymbol{X}\boldsymbol{w}^k - \boldsymbol{r}^{\mathrm{b}}\right)^+.$$

Plot the trade-off curve of regression error vs. sparsity level for each method (by varying the hyper-parameter λ).

Exercise 13.10: FDR-controlling method for sparse index tracking

Download price data corresponding to some financial index, such as the S&P 500, and the corresponding constituent N assets for some period of time. Then, construct the benchmark return vector $\boldsymbol{r}^{\mathrm{b}}$ and the assets' return matrix $\boldsymbol{X}$, and formulate the sparse index tracking problem

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda\|\boldsymbol{w}\|_0 \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1, \ \boldsymbol{w} \geq \mathbf{0}. \end{array}$$

a. Approximate the sparsity regularizer with the ℓ_1-norm:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda\|\boldsymbol{w}\|_1 \\ \text{subject to} & \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1, \ \boldsymbol{w} \geq \mathbf{0}. \end{array}$$

   Then solve the problem for different values of λ and plot the trade-off curve of regression error vs. sparsity level.

b. Employ the T-Rex method to automatically choose the active assets with FDR control.