Portfolio Optimization

Index Tracking Portfolios

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

portfoliooptimizationbook.com

Latest update: 2025-10-18

Outline

Executive Summary

- **Market efficiency debate**: Efficient-market hypothesis suggests prices reflect all public information, while opposing views argue markets are inefficient and irrational.
- **Active management reality**: ~95% of investment funds fail to beat the market despite claims from funds and financial experts.
- **Index tracking as alternative**: Exploring systematic approaches to market tracking instead of active strategies.
- These slides explore (Palomar 2025, chap. 13):
  - Heuristic and discretionary methods
  - Advanced optimization formulations
  - Recent techniques for automated asset selection with statistical controls

Outline

Active Versus Passive Strategies

Active Strategies

- Premise: market is not perfectly efficient (Shiller 1981).
- Goal: add value by choosing high-performing assets to beat market performance.
- Examples:
  - Mean–variance portfolios (Palomar 2025, chap. 7).
  - High-order portfolios (Palomar 2025, chap. 9).
  - Portfolios with alternative measures of risk (Palomar 2025, chap. 10).
  - Risk parity portfolios (Palomar 2025, chap. 11).
  - Graph-based portfolios (Palomar 2025, chap. 12).

Passive Strategies

- Assumption: market is efficient (Fama 1970).
- Belief: prices reflect all available information, making it impossible to beat the market in the long run (Malkiel 1973).
- Approach: avoid frequent trading to minimize fees and focus on infrequent rebalancing.
- Focus: index tracking as a method of passive investment (Prigent 2007; Benidis, Feng, and Palomar 2018a), (Palomar 2025, chap. 13).

Beating the Market

Consistency in Beating the Market

- Common perception: only winners are noticed; losers are often ignored.
- Importance: proper data analysis requires data from both winners and losers to distinguish luck from skill.

Individual Investors' Performance

- Individual investors trading stocks directly face significant performance penalties for active trading (B. M. Barber and Odean 2000).
- Question: can expert financial managers perform better?

Fund Managers' Performance

- Studies show 85-95% of actively managed mutual funds fail to outperform their benchmark indices over the long term (Malkiel 1973).
- Variability: performance can vary based on time period, benchmark index, and specific group of fund managers.

Beating the Market

Provocative Conclusions (Malkiel 1973)

- "Investors are far better off buying and holding an index fund than attempting to buy and sell individual securities or actively managed mutual funds."
- "The market prices stocks so efficiently that a blindfolded chimpanzee throwing darts at the stock listings can select a portfolio that performs as well as those managed by the experts."

Investor Realization

- Most investors recognize their limitations in stock-picking and trade management.
- Even most professionals struggle to outperform benchmarks.
- Result: paying high mutual fund expenses for underperformance is illogical.
- Trend: rise of indices and inexpensive exchange-traded funds (ETFs).

What Is a Financial Index?

Financial Index

- Definition: a collection of carefully selected assets to capture the value of a specific market or segment.
- Nature: equivalent to a hypothetical portfolio of assets; not a tradable financial instrument.

Index Composition

- Defined by the universe of assets and their composition percentages.
- **Capitalization-weighted (cap-weighted) index:**
  - assets weighted based on market capitalization (outstanding shares $\times$ share price)
  - index value proportional to the weighted average of the capitalization of underlying assets.
- **Other index construction methods:**
  - price-weighted
  - equal-weighted
  - fundamentally-weighted
  - factor-weighted.

What Is a Financial Index?

Example: S&P 500

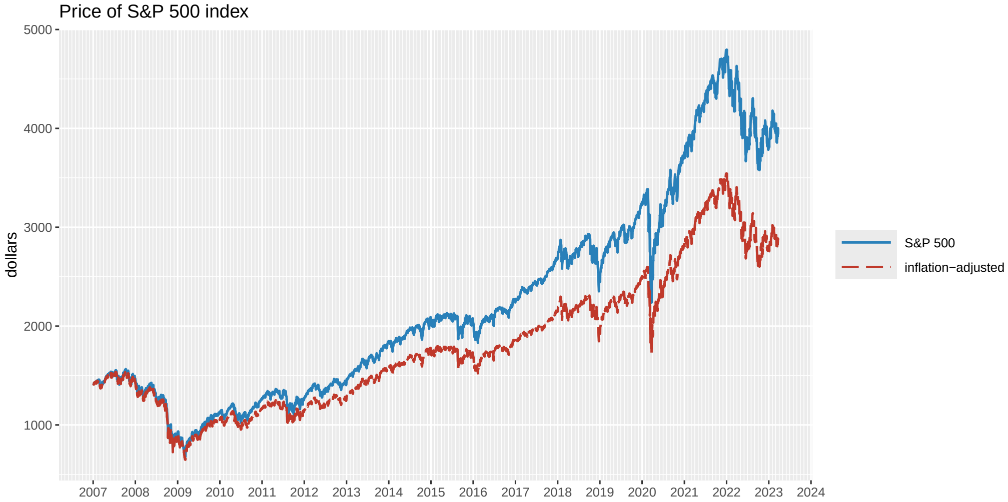
- One of the world's best-known cap-weighted indices.
- Commonly used benchmark for the U.S. stock market.
- Historical performance: next figure shows the S&P 500 price over more than a decade; prices have historically risen, providing reasonable returns simply by following the market; period includes the 2007–2008 global financial crisis and the COVID-19 recession.

Other Financial Indices

- Thousands of financial indices exist, covering various asset classes, sectors, and regions.
- Examples:
  - Dow Jones Industrial Average (USA)
  - Nasdaq Composite (USA)
  - FTSE 100 (UK)
  - Nikkei 225 (Japan)
  - DAX (Germany)
  - Hang Seng Index (Hong Kong)
  - Shanghai Shenzhen CSI 300 Index (mainland China)
  - IBEX 35 (Spain)

What Is a Financial Index?

Price evolution of the S&P 500 index:



Index Tracking

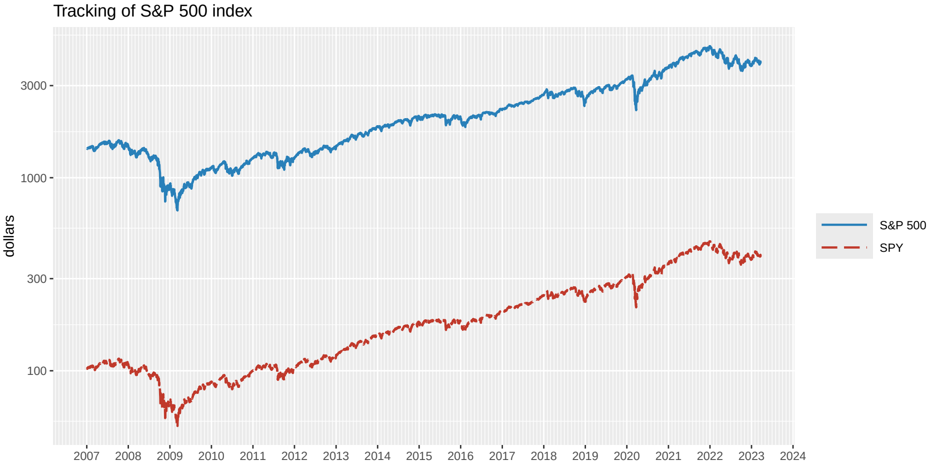
Index Tracking (Index Replication)

- Definition: passive portfolio management strategy to reproduce the performance of a market index.
- Practical implementation: construct a real portfolio that mimics the index as closely as possible.

Exchange-Traded Funds (ETFs)

- ETFs track indices and can be directly traded by investors.
- Example: over 250 ETFs track the S&P 500 index as of 2023.
- Most popular ETF: SPDR S&P 500 ETF (ticker SPY), the largest and oldest ETF in the world.
- Performance: next figure shows the S&P 500 index and SPDR S&P 500 ETF (SPY); SPY value is approximately $1/10$ of the S&P 500 level; ratio varies from $1/14$ to $1/10$ over time (2007–2023), but short-term variations are more critical for hedging.

Tracking of the S&P 500 index by the SPDR S&P 500 ETF:



Index Tracking

Full Replication

- Approach: buy appropriate quantities of all assets composing the index.
- Requirements: knowledge of precise index composition and regular rebalancing.
- Challenges: transaction costs and managing less liquid assets.

Sparse Index Tracking (Portfolio Compression)

- Approach: hold active positions on a reduced basket of representative assets.
- Example: instead of 500 assets for the S&P 500, invest in only 20 properly selected assets.
- Advantages: lower transaction costs and easier management.
- Related techniques: sparse regression techniques in statistics.
- References: (Jansen and Van Dijk 2002; Maringer and Oyewumi 2007; Scozzari et al. 2013; Xu, Lu, and Xu 2016; Benidis, Feng, and Palomar 2018b, 2018a; Machkour, Palomar, and Muma 2024).

Outline

Sparsity and Cardinality

Sparsity

- Definition: a vector is **sparse** if it has many elements equal to zero.
- Importance: sparsity is crucial in various problems where controlling the sparsity level is desired (Elad 2010).

Cardinality

- Definition: the **cardinality** of a vector $\boldsymbol{x} \in \mathbb{R}^N$, denoted by card($\boldsymbol{x}$), is the number of nonzero elements:

$$\text{card}(\boldsymbol{x}) \triangleq \sum_{i=1}^{N} 1\{x_i \neq 0\},$$

  where $1\{\cdot\}$ denotes the indicator function.
- Notation: often written as the ℓ_0-pseudo-norm $\|\boldsymbol{x}\|_0$, which is not a norm (not even convex).

Problem Formulation

Sparse Regression: Regression problem with extra requirement of solution being sparse.

Common Formulations

- **Regularized sparse least squares (LS):**

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda\|\boldsymbol{x}\|_0,$$

  where the parameter λ enforces more or less sparsity in the solution.
- **Constrained sparse LS:**

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

  where k controls the sparsity level.
- **Sparse underdetermined system of linear equations:**

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

  where the system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is underdetermined (having an infinite number of solutions).

Methods for Sparse Regression

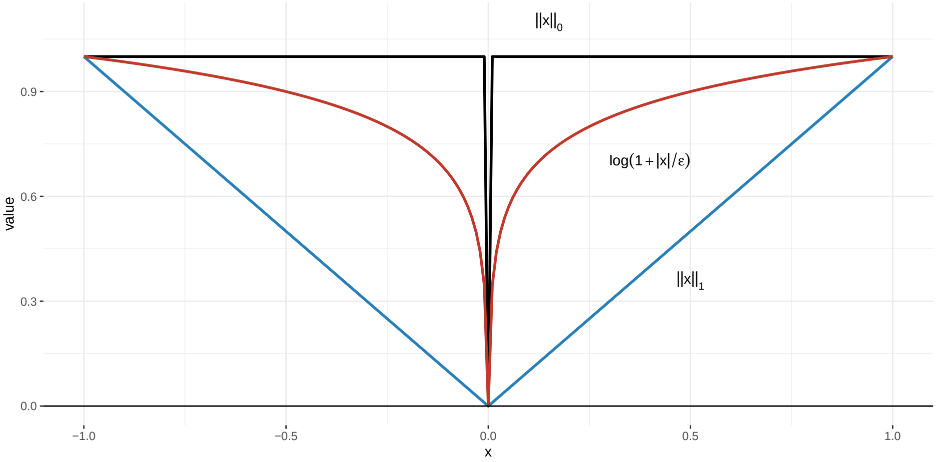
Challenges

- The cardinality operator is noncontinuous, nondifferentiable, and nonconvex.
- Developing practical algorithms under sparsity is complex.
- This has been a well-researched topic for decades, with mature approximate methods currently available.

Approximating the Cardinality Operator (Candès, Wakin, and Boyd 2008)

- **ℓ_1-norm approximation:** $\|\boldsymbol{x}\|_0$ is approximated by $\|\boldsymbol{x}\|_1$; univariate case: indicator function $1\{t \neq 0\}$ is approximated by the absolute value $|t|$.
- **Concave approximation:**
  - $\|\boldsymbol{x}\|_0$ is approximated by a concave function, making the sparse regression problem nonconvex
  - univariate case: indicator function $1\{t \neq 0\}$ is approximated by a concave function, such as the log-function $\log(1 + t/\varepsilon)$, where parameter ε controls the accuracy of the approximation.

Indicator function and approximations:



ℓ_1-Norm Approximation

LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani 1996)
Solves a convex quadratic problem using the ℓ_1-norm approximation:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \\ \text{subject to} & \|\boldsymbol{x}\|_1 \leq t, \end{array}$$

where the parameter t controls the sparsity level.

Elastic Net (Zou and Hastie 2005)

- Combines ℓ_1 and ℓ_2 penalties of LASSO and ridge methods.
- Addresses limitations of LASSO, especially with highly correlated variables.

Sparse Resolution of Underdetermined Systems (Candès and Tao 2005; Donoho 2006)
Can be recovered as a linear program under technical conditions:

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

Concave Approximation

Refined Concave Approximation

- When the ℓ_1-norm approximation is insufficient, a more refined concave approximation:

$$\|\boldsymbol{x}\|_0 \approx \sum_{i=1}^{N} \phi(|x_i|),$$

where $\phi(\cdot)$ is an appropriate concave function, such as the log-function $\phi(t) = \log(1 + t/\varepsilon)$ (Candès, Wakin, and Boyd 2008).

Nonconvex Problem Formulation

- Concave approximation leads to a nonconvex problem:

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

Solution Methods

- **Majorization-minimization framework:** used to derive iterative methods for problems with concave approximations of cardinality.
- **Iteratively reweighted least squares (IRLS) minimization:** another successful family of algorithms (Daubechies et al. 2010).

Majorization-Minimization (MM) Method Overview (Hunter and Lange 2004; Sun, Babu, and Palomar 2017) (Palomar 2025, Appendix B)

- MM simplifies complex optimization problems through iterative surrogate minimization.
- Iteratively produces a sequence $\boldsymbol{x}^0, \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots$ converging towards a solution $\boldsymbol{x}^\star$.

MM Iterative Process

- At each iteration k, MM uses a surrogate function $u(\boldsymbol{x}; \boldsymbol{x}^k)$ to approximate $f(\boldsymbol{x})$.
- The update rule is:

$$\boldsymbol{x}^{k+1} = \underset{\boldsymbol{x}\in\mathcal{X}}{\arg\min}\; u\left(\boldsymbol{x}; \boldsymbol{x}^k\right) \qquad k = 0, 1, 2, \ldots$$

Convergence Conditions for MM: Surrogate function $u(\boldsymbol{x}; \boldsymbol{x}^k)$ must satisfy:

- *Upper bound property*: $u(\boldsymbol{x}; \boldsymbol{x}^k) \geq f(\boldsymbol{x})$.
- *Touching property*: $u(\boldsymbol{x}^k; \boldsymbol{x}^k) = f(\boldsymbol{x}^k)$.
- *Tangent property*: differentiable with $\nabla u(\boldsymbol{x}; \boldsymbol{x}^k) = \nabla f(\boldsymbol{x})$.

Role of the Surrogate Function

- Acts as a majorizer, providing an upper bound to the original function.
- The name MM stems from the process of constructing and minimizing the majorizer.

Iterative Reweighted ℓ_1-Norm Minimization

Let's focus on the following sparse regression:

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \sum_{i=1}^{N} \log\left(1 + \frac{|x_i|}{\varepsilon}\right) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{array}$$

In order to use MM, we need a majorizer of the concave function:

Lemma: Majorizer of the log function

The concave function $\phi(t) = \log(1 + t/\varepsilon)$ is majorized at $t = t_0$ by its linearization:

$$\phi(t) \leq \phi(t_0) + \phi(t_0)'(t - t_0) = \phi(t_0) + \frac{1}{\varepsilon + t_0}(t - t_0).$$

Iterative Reweighted ℓ_1-Norm Minimization

The term $\log(1 + |x_i|/\varepsilon)$ is majorized at x_i^k by $\alpha_i^k |x_i|$ with weight $\alpha_i^k = \frac{1}{\varepsilon + |x_i^k|}$.

Iterative Algorithm
To solve the original nonconvex regression problem, solve the following sequence of convex problems for $k = 0, 1, 2, \ldots$

$$\begin{array}{ll} \underset{\boldsymbol{x}}{\text{minimize}} & \sum_{i=1}^{N} \alpha_i^k |x_i| \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{array}$$

where the objective function is a weighted ℓ_1-norm with weights $\alpha_i^k = 1/(\varepsilon + |x_i^k|)$.

Summary
The concave approximation of the sparse regression problem can be effectively solved by a sequence of iterative reweighted ℓ_1-norm minimization problems (Candès, Wakin, and Boyd 2008).

Outline

Sparse Index Tracking

Sparse Index Tracking
Sparse index tracking is an instance of sparse regression.

Challenge
Difficulty lies on the cardinality constraint or sparsity control.

Methods
Various methods proposed in the literature to address the challenge: (Jansen and Van Dijk 2002; Maringer and Oyewumi 2007; Scozzari et al. 2013; Xu, Lu, and Xu 2016; Benidis, Feng, and Palomar 2018b, 2018a).

Formulation Details
First, we elaborate on the choice of the tracking error function, as well on some details of the formulation such as fixed vs. time-varying portfolios, linear versus log-returns, plain versus cumulative returns.

Outline

Return of an Index or Benchmark

- Obtained from the returns of the N constituent assets $\boldsymbol{r}_t \in \mathbb{R}^N$ via the index portfolio $\boldsymbol{b}_t > \mathbf{0}$ (normalized to $\mathbf{1}^\mathsf{T}\boldsymbol{b}_t = 1$):

$$\boldsymbol{r}_t^\mathsf{T}\boldsymbol{b}_{t-1} = r_t^\mathrm{b}, \quad t = 1, \dots, T,$$

where vector $\boldsymbol{b}_t$ denotes the proportion of capital allocated to the assets (alternatively, it can be defined in terms of number of shares).

Fixed Portfolio Notation

- If the portfolio is fixed over time, $\boldsymbol{b}_t = \boldsymbol{b}$, the notation simplifies to:

$$\boldsymbol{X}\boldsymbol{b} = \boldsymbol{r}^\mathrm{b},$$

where matrix $\boldsymbol{X} \in \mathbb{R}^{T\times N}$ contains return vectors $\boldsymbol{r}_t$ along the rows, vector $\boldsymbol{r}^\mathrm{b} \in \mathbb{R}^T$ contains the returns of the index.

Designing a Sparse Portfolio

- Goal: design a sparse portfolio $\boldsymbol{w}_t$ such that $\boldsymbol{r}_t^\mathsf{T}\boldsymbol{w}_{t-1} \approx r_t^\mathrm{b}$.
- Simplified notation: assume a fixed portfolio over time $\boldsymbol{w}_t = \boldsymbol{w}$:

$$\boldsymbol{X}\boldsymbol{w} \approx \boldsymbol{r}^\mathrm{b}.$$

Tracking Error (TE)

- Simplest definition (Shapcott 1992; Jansen and Van Dijk 2002; Xu, Lu, and Xu 2016; Scozzari et al. 2013; Benidis, Feng, and Palomar 2018b, 2018a):

$$\mathrm{TE} = \frac{1}{T}\left\|\boldsymbol{r}^\mathrm{b} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2.$$

- Expansion with $\boldsymbol{X}\boldsymbol{b} = \boldsymbol{r}^\mathrm{b}$:

$$\mathrm{TE} = (\boldsymbol{b} - \boldsymbol{w})^\mathsf{T}\frac{1}{T}\boldsymbol{X}^\mathsf{T}\boldsymbol{X}(\boldsymbol{b} - \boldsymbol{w}).$$

- Alternative definition based on $\boldsymbol{b}$:

$$\mathrm{TE}^\mathrm{b} = (\boldsymbol{b} - \boldsymbol{w})^\mathsf{T}\boldsymbol{\Sigma}(\boldsymbol{b} - \boldsymbol{w}).$$

Fixed vs. Time-Varying Portfolio*

Traditional Tracking Error Definition

$$\mathrm{TE} = \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2.$$

- Convenient because it fits naturally in the context of sparse regression.
- However, index portfolio changes over time.
- Example: for cap-weighted indices, the normalized portfolio evolves as

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

Time-Varying Portfolio

- Approximating a time-varying portfolio $\boldsymbol{b}_t$ with a constant portfolio $\boldsymbol{w}$ seems nonsensical.
- In addition, a fixed portfolio $\boldsymbol{w}$ implies frequent rebalancing due to its time-varying nature:

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

Fixed vs. Time-Varying Portfolio*

Improving Tracking Error with Time-Varying Portfolios

- Approximation based on tracking the index $\boldsymbol{r}_t^{\mathsf{T}}\boldsymbol{w}_{t-1} \approx r_t^{\mathrm{b}}$:

$$\boldsymbol{w}_t \approx \frac{\boldsymbol{w}_{t-1} \odot (\boldsymbol{1} + \boldsymbol{r}_t)}{1 + r_t^{\mathrm{b}}} \approx \boldsymbol{w}_0 \odot \boldsymbol{\alpha}_t,$$

  where $\boldsymbol{w}_0$ is the initial portfolio and $\boldsymbol{\alpha}_t = \prod_{t'=1}^{t} \frac{\boldsymbol{1}+\boldsymbol{r}_{t'}}{1+r_{t'}^{\mathrm{b}}}$ denotes weights based on cumulative returns.

- Portfolio Return:

$$\boldsymbol{r}_t^{\mathsf{T}}\boldsymbol{w}_{t-1} \approx \boldsymbol{r}_t^{\mathsf{T}}\left(\boldsymbol{w}_0 \odot \boldsymbol{\alpha}_{t-1}\right) = \tilde{\boldsymbol{r}}_t^{\mathsf{T}}\boldsymbol{w}_0,$$

  where $\tilde{\boldsymbol{r}}_t = \boldsymbol{r}_t \odot \boldsymbol{\alpha}_{t-1}$ are properly weighted returns.

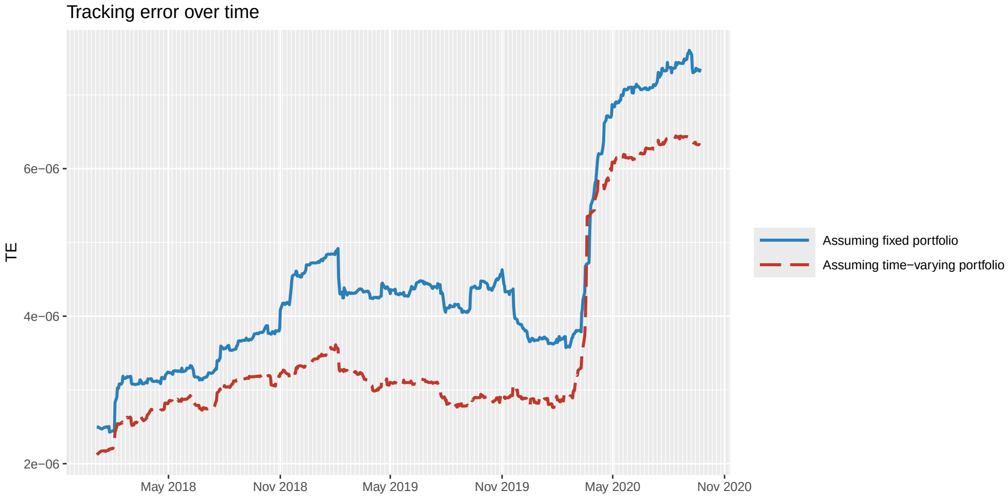
- Tracking Error for Time-Varying Portfolios:

$$\mathrm{TE}^{\text{time-varying}} = \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \tilde{\boldsymbol{X}}\boldsymbol{w}\right\|_2^2,$$

  where $\tilde{\boldsymbol{X}}$ contains the weighted returns $\tilde{\boldsymbol{r}}_t$ row-wise and $\boldsymbol{w}$ denotes the initial portfolio.

Fixed vs. Time-Varying Portfolio*

Tracking error over time of the S&P 500 index assuming fixed and time-varying portfolios:



Linear vs. Log-Returns*

Linear Returns

- Portfolio returns computed as $\boldsymbol{X}\boldsymbol{w}$ assume linear returns in $\boldsymbol{X}$ and $\boldsymbol{r}^{\mathrm{b}}$.
- Indeed, linear returns are additive along assets.

Log-Returns

- Log-returns $\boldsymbol{r}_t^{\log}$ and linear returns $\boldsymbol{r}^{\mathrm{lin}}$ are related:

$$\boldsymbol{r}_t^{\log} = \log\left(\mathbf{1} + \boldsymbol{r}_t^{\mathrm{lin}}\right) \approx \boldsymbol{r}_t^{\mathrm{lin}},$$

  due to $\log(1+x) \approx x$ for small x.
- Log-returns of the portfolio $\boldsymbol{w}$ can be approximated as:

$$\log\left(1 + \boldsymbol{w}^\mathsf{T}\boldsymbol{r}_t^{\mathrm{lin}}\right) \approx \boldsymbol{w}^\mathsf{T}\boldsymbol{r}_t^{\log}.$$

Practical Implication

- The difference between using linear or log-returns is negligible in practice.

Plain vs. Cumulative Returns

Minimizing Error in Period-by-Period Returns

- Implements a short-term index tracking (essential for hedging purposes against other investments).
- But does not imply better tracking of cumulative returns (or price) over time.
- Errors in returns can accumulate on a more positive or negative side.
- To control long-term deviation of cumulative returns, use long-term returns or cumulative returns (Benidis, Feng, and Palomar 2018a).

Cumulative Return or Price Error Measurement

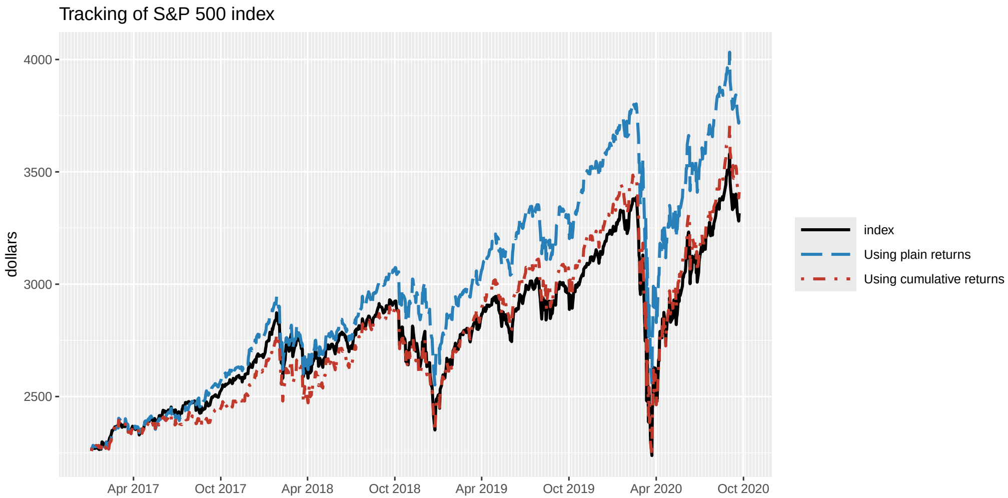
- Tracking error in terms of cumulative returns (Palomar 2025, chap. 13):

$$\mathrm{TE}^{\mathrm{cum}} = \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b,cum}} - \boldsymbol{X}^{\mathrm{cum}}\boldsymbol{w}\right\|_2^2,$$

  where the cumulative returns are $\boldsymbol{r}_t^{\mathrm{cum}} \approx \sum_{t'=1}^{t} \boldsymbol{r}_{t'}$.

Plain vs. Cumulative Returns

Tracking over time of the S&P 500 index using plain returns and cumulative returns:



Summary of Tracking Errors

Different data matrices employed in the definition of tracking error:

	tracking of returns	tracking of prices
assuming w fixed	$\boldsymbol{X}$	$\boldsymbol{X}^{\text{cum}}$
assuming w time-varying	$\tilde{\boldsymbol{X}}$	$\tilde{\boldsymbol{X}}^{\text{cum}}$

where

- $\boldsymbol{X}$: contains the plain returns $\boldsymbol{r}_t$ along the rows
- $\tilde{\boldsymbol{X}}$: contains the weighted plain returns $\tilde{\boldsymbol{r}}_t = \boldsymbol{r}_t \odot \boldsymbol{\alpha}_{t-1}$ along the rows
- $\boldsymbol{X}^{\text{cum}}$: contains the cumulative returns $\boldsymbol{r}_t^{\text{cum}} \approx \sum_{t'=1}^{t} \boldsymbol{r}_{t'}$ along the rows
- $\tilde{\boldsymbol{X}}^{\text{cum}}$: contains the weighted cumulative returns $\sum_{t'=1}^{t} \tilde{\boldsymbol{r}}_{t'}$.

Outline

Problem Formulation

Index Tracking Problem Formulation

- Formulated as a sparse regression problem:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 + \lambda\|\boldsymbol{w}\|_0 \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W}, \end{array}$$

  where:
  - parameter λ controls the level of sparsity in the portfolio,
  - constraint set $\mathcal{W}$ could be $\{\boldsymbol{w} \mid \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1, \boldsymbol{w} \geq \mathbf{0}\}$,
  - matrix $\boldsymbol{X}$ contains the returns of the assets.

Challenges with Cardinality Operator

- Cardinality operator $\|\boldsymbol{w}\|_0$ is noncontinuous, nondifferentiable, and nonconvex.
- Developing practical algorithms under sparsity is complex.
- Approach: start with heuristic methods and build up to state-of-the-art solutions based on MM.

Outline

Methods

Two-Step Approaches

- **Observation:** if active assets were known a priori, index tracking would be a trivial convex regression problem without sparsity control difficulty.
- **Two-step process:** first, select the active assets; then, compute the weights (Jansen and Van Dijk 2002).
- **Limitation:** not optimal; better results can be achieved by solving the problem jointly in a single step.

Joint Methods

- **Computationally intensive methods:** mixed integer programming, differential evolution techniques (Maringer and Oyewumi 2007).
- **Computationally feasible methods:** iterative reweighted ℓ_1-norm optimization method (Benidis, Feng, and Palomar 2018b, 2018a), projected gradient method (Xu, Lu, and Xu 2016).
- **Limitation:** these methods cannot theoretically guarantee a globally optimal solution.

Naive Two-Step Design

1. **Heuristic selection of active assets:** Select K active assets from a universe of N assets ($K \ll N$) (Jansen and Van Dijk 2002), based on different criteria:
   - Weight of the assets in the index definition (e.g., largest K assets in $\boldsymbol{b}$).
   - Market capitalization of the assets.
   - Strength of correlation between the assets and the index.

2. **Weight calculation:** Weights $\boldsymbol{w}$ proportional to the definition in $\boldsymbol{b}$ (scaled so that $\mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$):

$$\boldsymbol{w} = \frac{\boldsymbol{b} \odot \boldsymbol{s}}{\mathbf{1}^\mathsf{T}(\boldsymbol{b} \odot \boldsymbol{s})},$$

   where $\odot$ denotes Hadamard (elementwise) product and the Boolean pattern vector $\boldsymbol{s} \in \mathbb{R}^N$ is defined as

$$s_i = \begin{cases} 1 & \text{if the } i\text{-th asset is selected} \\ 0 & \text{otherwise.} \end{cases}$$

Two-Step Design with Refitted Weights

1. **Heuristic selection of active assets:** Select K active assets from a universe of N assets ($K \ll N$) as before and construct the Boolean pattern vector $\boldsymbol{s} \in \mathbb{R}^N$.

2. **Weight calculation:** Solve a simple regression problem over the active assets without having to deal with the sparsity issue:

$$\begin{array}{ll} \underset{\boldsymbol{w}}{\text{minimize}} & \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2 \\ \text{subject to} & \boldsymbol{w} \in \mathcal{W} \\ & \boldsymbol{w} \leq \boldsymbol{s}. \end{array}$$

Mixed Integer Programming Formulation

Mixed Integer Programming (MIP)

- **Nature:** variables constrained to a discrete set, making the problem nonconvex with exponential worst-case complexity.
- **Limitation:** impractical for large dimensionality (number of assets) (Scozzari et al. 2013).

MIP Formulation

- Reformulate the problem in terms of the Boolean pattern vector $\boldsymbol{s}$ (now a variable with nonconvex constraints!):

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

Evolutionary Algorithms

Evolutionary Algorithms

- **Definition:** optimization and search techniques inspired by natural evolution.
- **Operation:** work on populations of candidate solutions (individuals) that evolve over time to improve their fitness.
- **Process:** successful solutions are selected and combined to produce new candidate solutions, similar to natural selection.

Applications

- Optimization
- Machine learning
- Game playing
- Suitable for complex, multimodal, and noisy search spaces.

Advantages

- Ability to explore large solution spaces.
- Robustness against local optima.
- Potential for parallelization.

Evolutionary Algorithms

Popular Evolutionary Algorithms

- **Genetic algorithms:** use binary or symbolic representation for solutions; apply genetic operators like selection, crossover, and mutation.
- **Differential evolution:** focuses on continuous optimization problems; effective for high-dimensional, non-linear, and noisy problems.

Application in Index Tracking

- Employed to solve complicated nonconvex mixed-integer problem formulations.
- Examples: (Shapcott 1992), (Beasley, Meade, and Chang 2003), (Maringer and Oyewumi 2007).
- **Limitation:** high computational complexity due to the need to evolve a population of solutions over many generations to explore the nonconvex fitness surface.

Iterative Reweighted ℓ_1-Norm Method

ℓ_1-Norm Approximation

- Not suitable for portfolio optimization due to constraints $\mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$ and $\boldsymbol{w} \geq \mathbf{0}$, leading to $\|\boldsymbol{w}\|_1 = \mathbf{1}^\mathsf{T}\boldsymbol{w} = 1$.

Concave Approximation

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

MM Framework

- Address the nonconvex problem with the MM framework to obtain an iterative procedure called *iterative reweighted ℓ_1-norm method* (Benidis, Feng, and Palomar 2018b, 2018a).

Software Implementation

- R package `sparseIndexTracking` implements this method and extensions (Benidis and Palomar 2019).

MM Primer

Majorization-Minimization (MM) Method Overview (Hunter and Lange 2004; Sun, Babu, and Palomar 2017) (Palomar 2025, Appendix B)

- MM simplifies complex optimization problems through iterative surrogate minimization.
- Iteratively produces a sequence $\boldsymbol{x}^0, \boldsymbol{x}^1, \boldsymbol{x}^2, \ldots$ converging towards a solution $\boldsymbol{x}^\star$.

MM Iterative Process

- At each iteration k, MM uses a surrogate function $u(\boldsymbol{x}; \boldsymbol{x}^k)$ to approximate $f(\boldsymbol{x})$.
- The update rule is:

$$\boldsymbol{x}^{k+1} = \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{argmin}}\; u\left(\boldsymbol{x}; \boldsymbol{x}^k\right) \qquad k = 0, 1, 2, \ldots$$

Convergence Conditions for MM: Surrogate function $u(\boldsymbol{x}; \boldsymbol{x}^k)$ must satisfy:

- *Upper bound property*: $u(\boldsymbol{x}; \boldsymbol{x}^k) \geq f(\boldsymbol{x})$.
- *Touching property*: $u(\boldsymbol{x}^k; \boldsymbol{x}^k) = f(\boldsymbol{x}^k)$.
- *Tangent property*: differentiable with $\nabla u(\boldsymbol{x}; \boldsymbol{x}^k) = \nabla f(\boldsymbol{x})$.

Role of the Surrogate Function

- Acts as a majorizer, providing an upper bound to the original function.
- The name MM stems from the process of constructing and minimizing the majorizer.

Iterative Reweighted ℓ_1-Norm Method

Algorithm: Iterative reweighted ℓ_1-norm method for sparse index tracking

Initialization:

- Choose initial point $\mathbf{w}^0 \in \mathcal{W}$ and set iteration counter $k \leftarrow 0$.

Repeat (kth iteration):

1. Compute weights: $\alpha_i^k = \frac{1}{\varepsilon + \left| w_i^k \right|}$.
2. Solve the following weighted ℓ_1-norm problem to obtain $\mathbf{w}^{k+1}$:

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

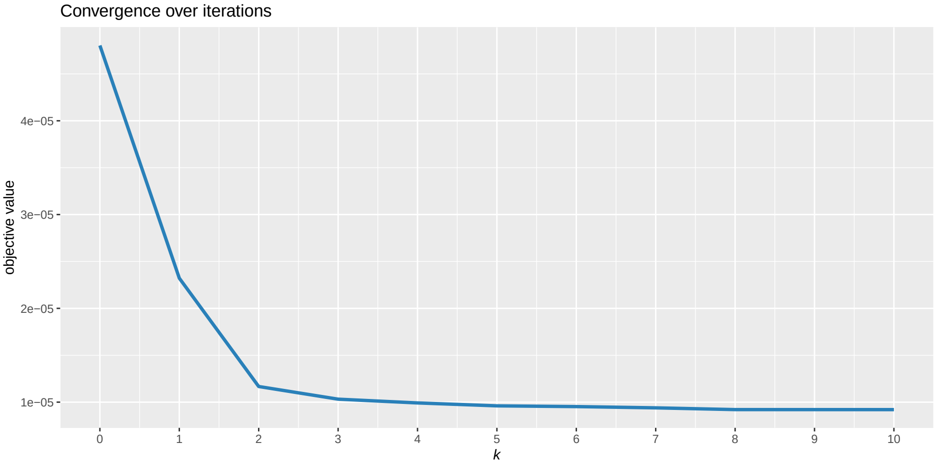
3. Increment the iteration counter: $k \leftarrow k + 1$.

Until: convergence

Outline

Convergence of the Iterative Reweighted ℓ_1-Norm Method

Convergence of the iterative reweighted ℓ_1-norm method for sparse index tracking:



Comparison of Algorithms

Comparison of Following Tracking Methods

- Naive two-step approach with proportional weights to the index definition.
- Two-step approach with refitted weights.
- Iterative reweighted ℓ_1-norm method.
- MIP formulation (excluded due to high computational complexity).

Figure Analysis

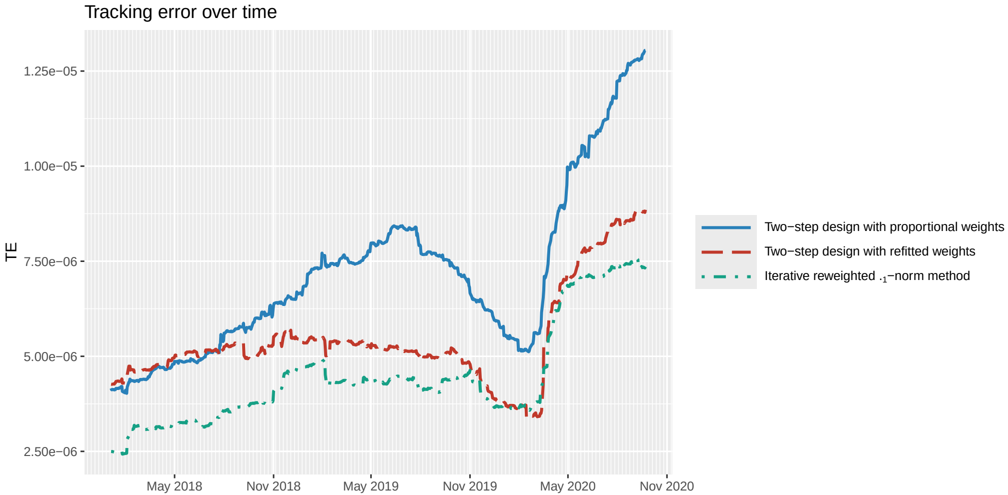
- **Tracking error over time:** shows tracking error over time for S&P 500 with $K = 20$ active assets; computed on a rolling window basis (two-year lookback), recomputed every six months.
- **Tracking error vs. active assets:** joint designs are superior to traditional two-step approaches; MIP formulation is impractical due to high computational cost; iterative reweighted ℓ_1-norm method exhibits low complexity, making it suitable in practice.

Conclusion

- Joint designs outperform traditional two-step approaches.
- Iterative reweighted ℓ_1-norm method is practical and efficient.
- Exhaustive numerical comparisons: (Benidis, Feng, and Palomar 2018b, 2018a).

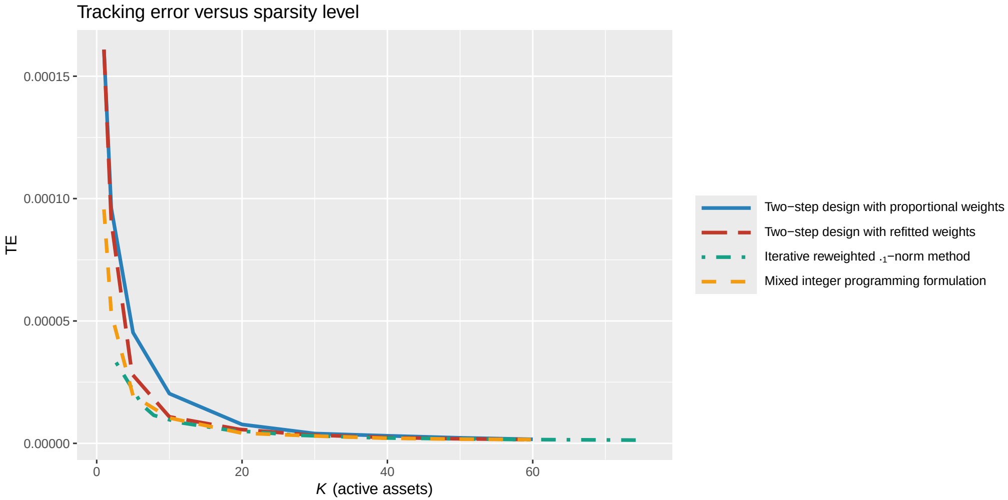
Comparison of Algorithms

Tracking error over time of the S&P 500 index for different algorithms:



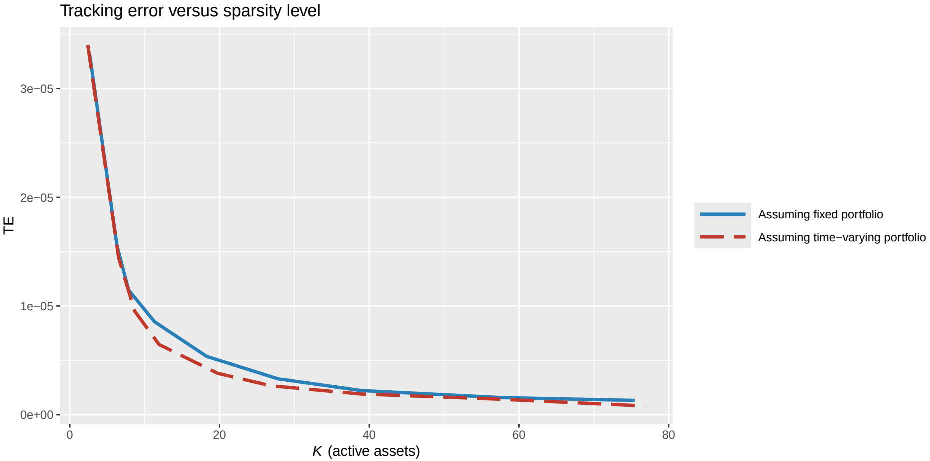
Comparison of Algorithms

Tracking error of the S&P 500 index versus active assets for different algorithms:



Comparison of Formulations

Tracking error of the S&P 500 index versus active assets assuming fixed and time-varying portfolios:



Outline

Enhanced Index Tracking

Enhanced Index Tracking: Variations on the basic index tracking formulation. Examples:

- increase returns via index-like portfolios with tactical tilts toward specific styles or stocks
- use alternative tracking error measures
- include group sparsity or holding constraints

Traditional Tracking Error

Based on the ℓ_2-norm between achieved returns $\boldsymbol{X}\boldsymbol{w}$ and benchmark returns $\boldsymbol{r}^{\mathrm{b}}$:

$$\mathrm{TE}(\boldsymbol{w}) = \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2.$$

Alternative Error Measures

- Change the norm (e.g., ℓ_1-norm or ℓ_p-norm (Beasley, Meade, and Chang 2003)).
- Change the error measure, such as the downside risk or the excess return $\frac{1}{T}\mathbf{1}^{\mathsf{T}}\left(\boldsymbol{X}\boldsymbol{w} - \boldsymbol{r}^{\mathrm{b}}\right)$ (Beasley, Meade, and Chang 2003; Dose and Cincotti 2005).
- Make the error measure robust to outliers.

Illustration: We next consider downside risk, the ℓ_1-norm tracking error, and a Huberized robust tracking error measure.

Downside Risk Error Measure

Downside Risk: Only considers when achieved returns are worse than the benchmark:

$$\mathrm{DR}(\boldsymbol{w}) = \frac{1}{T}\left\|\left(\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right)^{+}\right\|_2^2,$$

where $(\cdot)^{+} \triangleq \max(0, \cdot)$.
Convexity: $\mathrm{DR}(\boldsymbol{w})$ is a convex function of $\boldsymbol{w}$ and could be optimized with a solver.
Majorization: Alternatively, it can be majorized for the MM method using the ℓ_2-norm.

Lemma: Majorizer of the downside risk

The downside risk function $\mathrm{DR}(\boldsymbol{w})$ is majorized at $\boldsymbol{w} = \boldsymbol{w}_0$ by $\mathrm{TE}(\boldsymbol{w})$ with shifted benchmark returns $\tilde{\boldsymbol{r}}^{\mathrm{b}}$:

$$\mathrm{DR}(\boldsymbol{w}) \leq \frac{1}{T}\left\|\tilde{\boldsymbol{r}}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_2^2,$$

where $\tilde{\boldsymbol{r}}^{\mathrm{b}} = \boldsymbol{r}^{\mathrm{b}} + \left(\boldsymbol{X}\boldsymbol{w}_0 - \boldsymbol{r}^{\mathrm{b}}\right)^{+}$ (Benidis, Feng, and Palomar 2018b, 2018a).

Downside Risk Error Measure

Practical Implication

- We can now iteratively use the ℓ_2-norm based tracking error
$$\frac{1}{T}\left\|\tilde{\boldsymbol{r}}^{\mathrm{b}}-\boldsymbol{X}\boldsymbol{w}\right\|_2^2$$
using all the techniques and methods we have developed.

Interpretation

- The shifted benchmark returns
$$\tilde{\boldsymbol{r}}^{\mathrm{b}}=\boldsymbol{r}^{\mathrm{b}}+\left(\boldsymbol{X}\boldsymbol{w}_0-\boldsymbol{r}^{\mathrm{b}}\right)^{+}$$
are an improvement over the original returns $\boldsymbol{r}^{\mathrm{b}}$ for those returns that were outperformed by the nominal portfolio $\boldsymbol{w}_0$.

Algorithm Modification

- Index tracking under downside risk can be accomplished with the original algorithm with a slight modification: at each iteration k, use shifted benchmark returns instead of $\boldsymbol{r}^{\mathrm{b}}$:
$$\left(\tilde{\boldsymbol{r}}^{\mathrm{b}}\right)^{k}=\boldsymbol{r}^{\mathrm{b}}+\left(\boldsymbol{X}\boldsymbol{w}^{k}-\boldsymbol{r}^{\mathrm{b}}\right)^{+}.$$

ℓ_1-Norm Tracking Error

ℓ_1-Norm TE:

$$\mathrm{TE}_1(\boldsymbol{w}) = \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_1$$

Convexity: $\mathrm{TE}_1(\boldsymbol{w})$ is a convex function of $\boldsymbol{w}$ and could be optimized with a solver.

Majorization: Alternatively, it can be majorized for the MM method using the ℓ_2-norm.

Lemma: Majorizer of the ℓ_1-norm TE

The ℓ_1-norm TE function $\mathrm{TE}_1(\boldsymbol{w})$ is majorized at $\boldsymbol{w} = \boldsymbol{w}_0$ by a weighted version of the $\mathrm{TE}(\boldsymbol{w})$:

$$\mathrm{TE}_1(\boldsymbol{w}) \leq \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_{2,\boldsymbol{\alpha}}^2,$$

where $\|\boldsymbol{x}\|_{2,\boldsymbol{\alpha}}^2 \triangleq \sum_{i=1}^{T} \alpha_i x_i^2$ is the squared weighted ℓ_2-norm with weights $\boldsymbol{\alpha} = 1/(2|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}_0|)$ (Palomar 2025).

ℓ_1-Norm Tracking Error

Practical Implication

- We can now iteratively use the ℓ_2-norm based tracking error

$$\frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}}-\boldsymbol{X}\boldsymbol{w}\right\|_{2,\boldsymbol{\alpha}}^{2}$$

  using all the techniques and methods we have developed.

Interpretation

- Weights $\boldsymbol{\alpha}=1/(2|\boldsymbol{r}^{\mathrm{b}}-\boldsymbol{X}\boldsymbol{w}_0|)$ in the weighted ℓ_2-norm TE down-weight errors to grow linearly like in the ℓ_1-norm.

Algorithm Modification

- Index tracking under the ℓ_1-norm TE can be accomplished with the original algorithm with a slight modification: at each iteration k, use the weighted ℓ_2-norm with weights:

$$\boldsymbol{\alpha}^{k}=\frac{1}{2|\boldsymbol{r}^{\mathrm{b}}-\boldsymbol{X}\boldsymbol{w}^{k}|}.$$

Robust Tracking Error Measures

Robustness Against Outliers

- ℓ_2-norm is sensitive to large errors due to squaring.
- ℓ_1-norm is naturally robust.

Huber Penalty Function

- Robust version of the ℓ_2-norm:

$$\phi^{\text{hub}}(x) = \begin{cases} x^2, & |x| \leq M \\ M\left(2|x| - M\right) & |x| > M \end{cases}$$

- Behavior: Square function for $|x| \leq M$, linear function for $|x| > M$.

Huberized Tracking Error

$$\text{Hub-TE}(\boldsymbol{w}) = \frac{1}{T}\sum_{t=1}^{T} \phi^{\text{hub}}(r_t^{\text{b}} - \boldsymbol{X}_{t,:}\boldsymbol{w}).$$

Robust Tracking Error Measures

Convexity: Hub-TE($\boldsymbol{w}$) is a convex function of $\boldsymbol{w}$ and could be optimized with a solver.

Majorization: Alternatively, it can be majorized for the MM method.

Lemma: Majorizer of the Huberized TE

The Huberized TE function Hub-TE($\boldsymbol{w}$) is majorized at $\boldsymbol{w} = \boldsymbol{w}_0$ by a weighted version of the TE($\boldsymbol{w}$):

$$\text{Hub-TE}(\boldsymbol{w}) \leq \frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_{2,\boldsymbol{\alpha}}^{2} + \text{const},$$

where $\|\boldsymbol{x}\|_{2,\boldsymbol{\alpha}}^2 \triangleq \sum_{i=1}^{T} \alpha_i x_i^2$ is the squared weighted ℓ_2-norm with weights

$$\boldsymbol{\alpha} = \min\left(\mathbf{1}, \frac{M}{|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}_0|}\right)$$

(Benidis, Feng, and Palomar 2018b, 2018a).

Robust Tracking Error Measures

Practical Implication

- We can now iteratively use the ℓ_2-norm based tracking error

$$\frac{1}{T}\left\|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}\right\|_{2,\boldsymbol{\alpha}}^{2}$$

  using all the techniques and methods we have developed.

Interpretation

- Weights $\boldsymbol{\alpha} = \min\left(\mathbf{1}, \frac{M}{|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}_0|}\right)$ in the weighted ℓ_2-norm TE down-weight errors larger than M so squared values grow linearly.

Algorithm Modification

- Index tracking under the Huberized TE can be accomplished with the original algorithm with a slight modification: at each iteration k, use the weighted ℓ_2-norm with weights:

$$\boldsymbol{\alpha}^{k} = \min\left(\mathbf{1}, \frac{M}{|\boldsymbol{r}^{\mathrm{b}} - \boldsymbol{X}\boldsymbol{w}^{k}|}\right).$$

Numerical Experiments

Comparison of Tracking Error Measures for S&P 500 Index

- Basic TE
- Huberized TE
- ℓ_1-norm TE
- DR

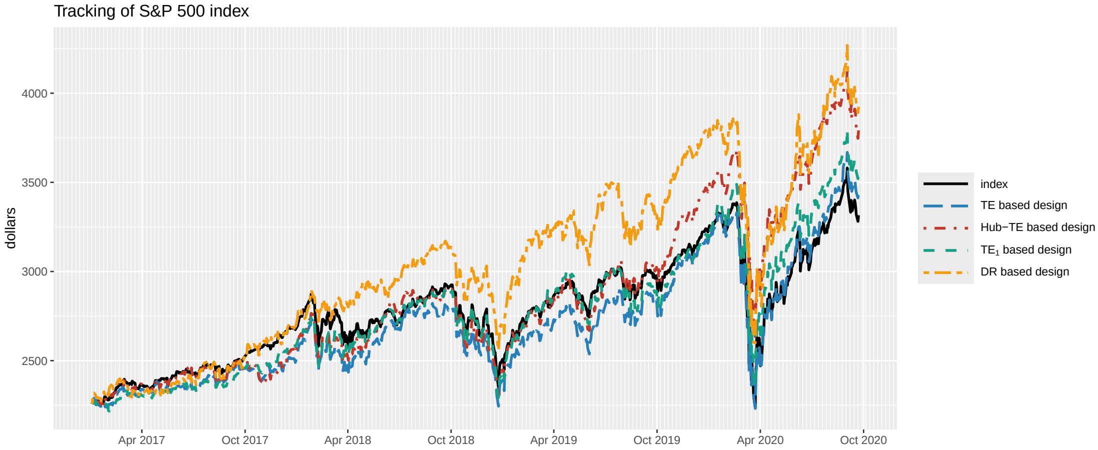
Numerical Experiments

- Shows tracking over time with approximately $K = 20$ active assets.
- Tracking portfolios computed on a rolling window basis with a two-year lookback period, recomputed every six months.

Observations

- Design based on downside risk (DR) beats the market, suitable for investment purposes.
- Other measures (TE, Huberized TE, ℓ_1-norm TE) generally track the index in both directions, appropriate for hedging purposes.

Tracking over time of the S&P 500 index under different tracking error measures:



Outline

Automatic Sparsity Control

Sparse Index Tracking Formulation

- **Objective:** includes regularization term $\lambda\|\boldsymbol{w}\|_0$, where λ is a hyper-parameter for desired sparsity.
- **Alternative formulation:** move sparsity term to constraints as $\|\boldsymbol{w}\|_0 \leq k$, with k denoting desired sparsity level.
- **Trade-off:** adjusting λ or k achieves different points on error vs. sparsity trade-off curves.

Choosing an Operating Point on the Trade-Off Curve

- **Goal:** select a proper point on the error versus sparsity trade-off curve without computing the entire curve.
- **Existing tuning approaches:**
  - **cross-validation:** use cross-validation to evaluate performance for different values of λ or k and select the one that minimizes the validation error;
  - **analytical methods:** employ automated hyper-parameter tuning methods such as Bayesian optimization or random search to efficiently explore the hyper-parameter space.

False Discovery Rate (FDR)

Choosing an Operating Point on the Trade-Off Curve

- Use concepts from statistics and hypothesis testing.
- **False discovery rate (FDR):** probability of wrongly including a variable in a regression problem.
- **Applications:** in genomics, including wrong variables can be catastrophic; in finance, many claimed research findings may be false (Harvey, Liu, and Zhu 2016).

Controlling FDR

- **Ideal method:** decide whether to include a variable by controlling the FDR.
- **Low-dimensional problems:** seminal paper in 1995 (Benjamini and Hochberg 1995).
- **High-dimensional problems:**
  - **knockoffs method:** fictitious variables mimicking the covariance structure of original variables, but high computational cost (R. F. Barber and Candès 2015);
  - **T-Rex (Terminating-Random EXperiments):** based on dummies, orders of magnitude lower computational cost (Machkour, Muma, and Palomar 2022).

Practical Approach

- Use FDR-controlling methods like T-Rex to determine the optimal sparsity level.

FDR for Index Tracking

FDR Control in Sparse Index Tracking

- **Robust approach:** control the FDR instead of selecting sparsity level through trial and error.
- **Interpretation of FDR:** all assets in an index are valid variables; many assets become redundant due to high correlation with selected assets; selecting these redundant assets is considered a "false discovery."

T-Rex Method for Sparse Index Tracking

- **Application:** developed in (Machkour, Palomar, and Muma 2024).
- **Advantage:** automatically selects assets by controlling the FDR.
- **Implementation:** R package `TRexSelector` (Machkour et al. 2022) implements the T-Rex method.

Practical Benefit

- Avoids the need to fix or tune the hyper-parameter λ in the sparse regression formulation.
- Ensures a more precise and robust selection of assets.

Numerical Experiments

Comparison of Tracking Portfolios

- Standard sparse penalized regression formulation.
- FDR-controlling T-Rex method (Machkour, Palomar, and Muma 2024).

Numerical Experiments

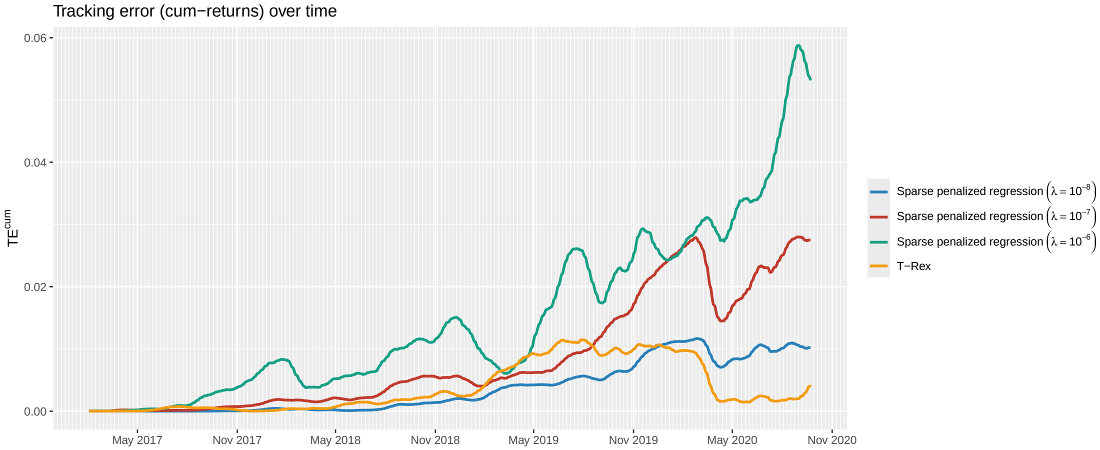
- Tracking portfolios computed on a rolling window basis with a two-year lookback period, recomputed every six months.
- Tracking error over time in terms of period-by-period returns and cumulative returns.
- Cardinality of the portfolios over time.

Observations

- Standard sparse penalized regression formulation is sensitive to the choice of parameter λ, resulting in varying tracking error and cardinality.
- FDR-controlling T-Rex method automatically selects the appropriate sparsity level without parameter tuning.
- Computational cost of T-Rex is slightly higher than solving the standard sparse penalized regression formulation for a fixed λ but lower than solving it for a range of λ values.

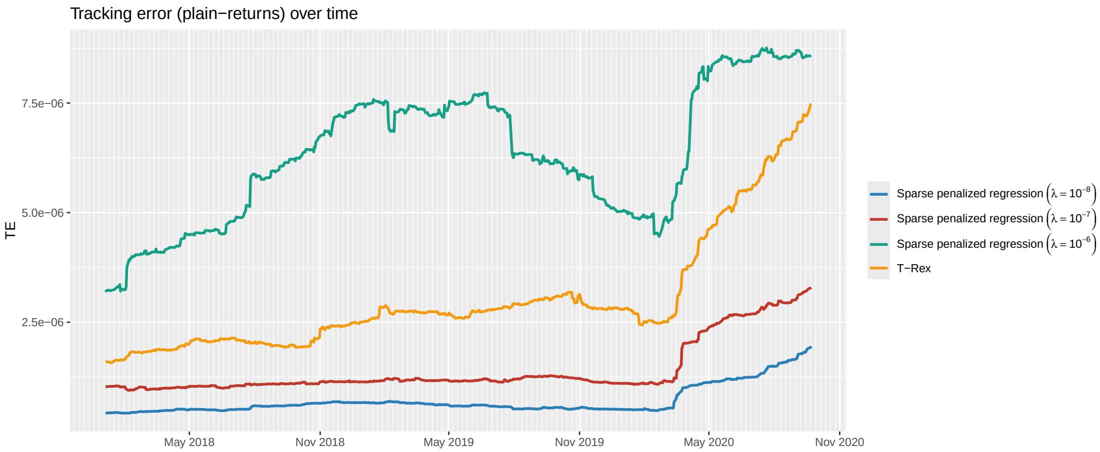
Conclusion: T-Rex method provides a robust and automatic way to control sparsity.

Tracking over time of the S&P 500 index with and without FDR control:



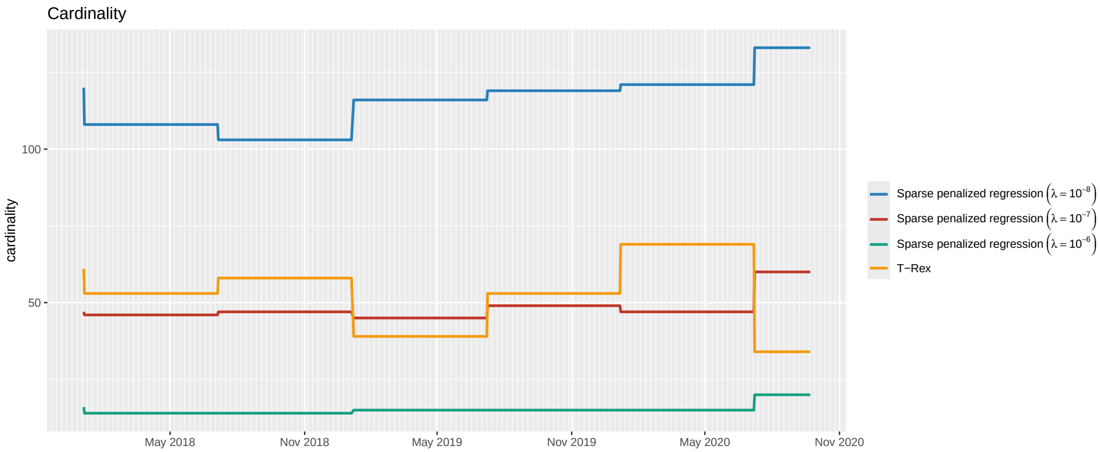
Numerical Experiments

Tracking over time of the S&P 500 index with and without FDR control:



Numerical Experiments

Tracking over time of the S&P 500 index with and without FDR control:



Outline

Summary

Active and passive strategies are crucial in financial investment, each justified by different market views. Key points include:

- **Passive investing** avoids fees and poor performance from frequent active trading.
- **Index tracking** mimics an index, assuming market efficiency and that it can't be beaten.
- **Financial indices** number in the thousands, covering various asset classes, sectors, and regions (e.g., S&P 500), with numerous ETFs available for direct trading.
- **Sparse index tracking** approximates an index using few active assets, involving a tracking error measure and sparsity control.
- **Tracking error measures** include the ℓ_2-norm, downside risk, ℓ_1-norm, and Huberized robust versions.
- **Algorithms** for index tracking, like the iterative reweighted ℓ_1-norm method, balance tracking error and sparsity with low computational cost.
- **Sparsity level** is often set by trial and error, but the new **FDR-controlling index tracking method** automatically determines it using hypothesis testing.

Barber, B. M., and T. Odean. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *The Journal of Finance* 55 (2): 773–806.

Barber, R. F., and E. J. Candès. 2015. "Controlling the False Discovery Rate via Knockoffs." *The Annals of Statistics* 43 (5): 2055–85.

Beasley, J. E., N. Meade, and T. J. Chang. 2003. "An Evolutionary Heuristic for the Index Tracking Problem." *European Journal of Operational Research* 148: 621–43.

Benidis, K., Y. Feng, and D. P. Palomar. 2018a. *Optimization Methods for Financial Index Tracking: From Theory to Practice*. Foundations and Trends in Optimization, Now Publishers.

———. 2018b. "Sparse Portfolios for High-Dimensional Financial Index Tracking." *IEEE Transactions on Signal Processing* 66 (1): 155–70.

References II

Benidis, K., and D. P. Palomar. 2019. *sparseIndexTracking: Design of Portfolio of Stocks to Track an Index*. https://CRAN.R-project.org/package=sparseIndexTracking.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 57 (1): 289–300.

Candès, E. J., and T. Tao. 2005. "Decoding by Linear Programming." *IEEE Transactions on Information Theory* 51 (12): 4203–15.

Candès, E. J., M. B. Wakin, and S. P. Boyd. 2008. "Enhancing Sparsity by Reweighted ℓ_1 Minimization." *Journal of Fourier Analysis and Applications* 14: 877–905.

Daubechies, I., R. Devore, M. Fornasier, and C. S. N. Güntürk. 2010. "Iteratively Reweighted Least Squares Minimization for Sparse Recovery." *Communications on Pure and Applied Mathematics* 63: 1–38.

Donoho, D. L. 2006. "Compressed Sensing." *IEEE Transactions on Information Theory* 52 (4): 1289–1306.

Dose, C., and S. Cincotti. 2005. "Clustering of Financial Time Series with Application to Index and Enhanced Index Tracking Portfolio." *Physica A: Statistical Mechanics and Its Applications* 355 (1): 145–51.

Elad, M. 2010. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer.

Fama, E. F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *The Journal of Finance* 25 (2): 383–417.

Harvey, C. R., Y. Liu, and H. Zhu. 2016"… And the Cross-Section of Expected Returns." *Review of Financial Studies* 29 (1): 5–68.

Hunter, D. R., and K. Lange. 2004. "A Tutorial on MM Algorithms." *The American Statistician* 58: 30–37.

Jansen, R., and R. Van Dijk. 2002. "Optimal Benchmark Tracking with Small Portfolios." *Journal of Portfolio Management* 28 (2): 33–39.

Machkour, J., M. Muma, and D. P. Palomar. 2022. "The Terminating-Random Experiments Selector: Fast High-Dimensional Variable Selection with False Discovery Rate Control." *Available at arXiv*. https://doi.org/10.48550/arXiv.2110.06048.

Machkour, J., D. P. Palomar, and M. Muma. 2024. "FDR-Controlled Portfolio Optimization for Sparse Financial Index Tracking." *Available at arXiv*. https://doi.org/10.48550/arXiv.2401.15139.

References V

Machkour, J., S. Tien, D. P. Palomar, and M. Muma. 2022. *TRexSelector: T-Rex Selector: High-Dimensional Variable Selection & FDR Control*. https://CRAN.R-project.org/package=TRexSelector.

Malkiel, B. G. 1973. *A Random Walk down Wall Street*. New York: W. W. Norton.

Maringer, D., and O. Oyewumi. 2007. "Index Tracking with Constrained Portfolios." *Intelligent Systems in Accounting, Finance and Management* 15 (1-2): 57–71.

Palomar, D. P. 2025. *Portfolio Optimization: Theory and Application*. Cambridge University Press.

Prigent, J. L. 2007. *Portfolio Optimization and Performance Analysis*. CRC Press.

Scozzari, A., F. Tardella, S. Paterlini, and T. Krink. 2013. "Exact and Heuristic Approaches for the Index Tracking Problem with UCITS Constraints." *Annals of Operations Research* 205 (1): 235–50.

References VI

Shapcott, J. 1992. "Index Tracking: Genetic Algorithms for Investment Portfolio Selection." Report EPCC-SS92-24, Edinburgh Parallel Computing Centre, The University of Edinburgh.

Shiller, R. J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3): 421–36.

Sun, Y., P. Babu, and D. P. Palomar. 2017. "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning." *IEEE Transactions on Signal Processing* 65 (3): 794–816.

Tibshirani, R. 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 58 (1): 267–88.

References VII

Xu, F., Z. Lu, and Z. Xu. 2016. "An Efficient Optimization Approach for a Cardinality-Constrained Index Tracking Problem." *Optimization Methods and Software* 31 (2): 258–71.

Zou, H., and T. Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 67 (2): 301–20.