ML in Finances

Пункт 1

На лекции мы показали, что задача Index tracking сводится к

$$\min_w \operatorname{Var}\left[R_B - w'R\right] = \min_w \operatorname{Var}[w^{'}Z] = \min_w \operatorname{E}[[Z^{(N)} - w'_{N-1}Z*]^2],$$

то есть, к задаче МНК. При этом, добавляя L2 регуляризацию, (в своем алгоритме использую L2 потому что алгоритм Франка Вульфа, который я буду использовать, подразумевает использование всех assets), получаем:

$$w^*_{-N} = \min_w \operatorname{E}[[Z^{(N)} - w'_{N-1}Z*]^2] + \lambda||w_{-N}||_2$$

не забываем про условия на портфель (на слайдах решается безусловная задача):

$$s.t\ w_i \geq 0, \sum_i w_i = 1$$

Сформулируем задачу LMO(c) $= \min\limits_{s\in\chi}(s, c)$:

Л1

Покажите, что для симплекса:

$$\mathcal{X} = \left\{s \in \mathbb{R}^d \mid s \succeq 0,\ \mathbf{1}^\top s = R\right\}$$

решение LMO:

$$s^* = Re_i,\ \text{где } i = \arg\min_{j=\overline{1,d}} c_j.$$

Доказательство. Запишем лагранжиан:

$$\mathcal{L}(s, \lambda, \nu) = \langle s, c\rangle - \lambda^\top s - \nu\left(1^\top s - R\right).$$

Воспользуемся теоремой Каруша–Куна–Таккера:

1. Ограничения: $s \succeq 0,\ 1^\top s = R$; 2. Неотрицательность: $\lambda \succeq 0$; 3. Дополняющая нежёсткость: $\lambda_i s_i = 0$; 4. Стационарность: $c - \lambda - \nu = 0$.

Возьмём

$$s^* = Re_i, \quad \lambda_j = c_j - c_i, \quad \nu = -c_i,$$

где

$$i = \operatorname{argmin}_{j=1,\ldots,d} c_j.$$

Алгоритм Франка–Вульфа

Для дуального зазора на k-й итерации используем

$$\mathrm{gap}(x^k) = \big\langle \nabla f(x^k),\, x^k - s^k \big\rangle.$$

Усреднённое значение метрики $\mathrm{gap}(x^k)$ задаётся как

$$\mathrm{gap} = \frac{1}{k}\sum_{t=0}^{k} \mathrm{gap}(x^t).$$

Псевдокод алгоритма

Инициализация:

- шаг $\{\gamma_k = \frac{2}{k+2}\}_{k=0}^{K-1}$;
- начальная точка $x^0 \in \mathbb{R}^d$;
- максимальное число итераций K.

k-я итерация ($k = 0, 1, \ldots, K-1$):

1. найти направление
$$s^k \in \arg\min_{s\in\mathcal{X}} \big\langle s,\, \nabla f(x^k)\big\rangle;$$
2. обновить итерат
$$x^{k+1} = (1-\gamma_k)x^k + \gamma_k s^k.$$

Условие остановки:

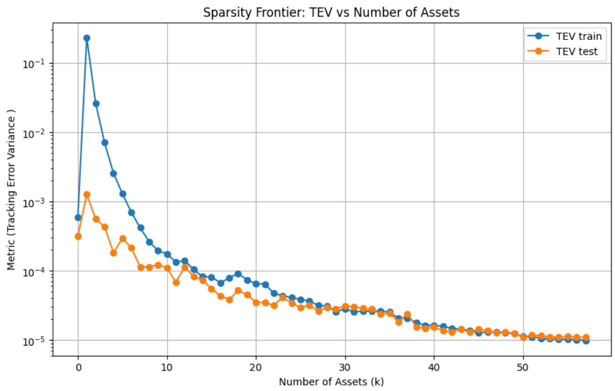
- достигнуто максимальное число итераций K; или
- $\mathrm{gap}(x^k) < \varepsilon$.

Выход: итоговая точка x^K.

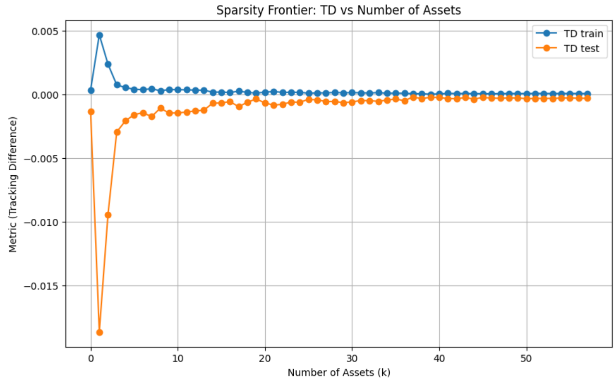
Коэффициент регуляризации λ возьмем L $= \lambda_{max}(A_{train}A_{train}^T)$, $\lambda_{max}(A)$ - максимальное собственное значение матрицы A (эмпирический результат), learning rate обычно берется $\gamma_k = \frac{2}{k+2}$, начальная точка $x_0 = \overline{0}$. В качестве обучающей выборки беру первые 8 лет, в качестве тестовой последние 2 года. Этапа подбора гиперпараметров у меня нет.

Также нетрудно доказать, что алгоритм Франка-Вульфа на k-ом шаге выдаст решение, в котором не более $k+1$ ненулевой компоненты (при условии $x_0 = \overline{0}$). При этом на каждой итерации решение достаточно хорошее (не оптимальное, потому что, по сути, алгоритм добавляет векторы жадно, но достаточно хорошее, что видно ниже), поэтому выбор алгоритма Франка-Вульфа, в том числе, обусловлен этой особенностью, которая нужна для анализа $Sparsity$.

Пункт 2



(a) Tracking Error Variance (TEV)



(b) Tracking Difference (TD)

Рис. 1: Зависимость метрик качества от разреженности портфеля (количества активов k). Графики показывают сходимость ошибки на тренировочной и тестовой выборках.

Таблица 1: Summary Statistics of Weekly Returns: FTSE 100 vs Tracking Portfolio

Series	Mean	Median	Std Dev	Variance	Min	Max
FTSE 100 Index	0.000538	0.001946	0.023225	0.000539	-0.210469	0.134092
Tracking Portfolio	0.001127	0.002247	0.023033	0.000531	-0.205758	0.129122

Пункт 3

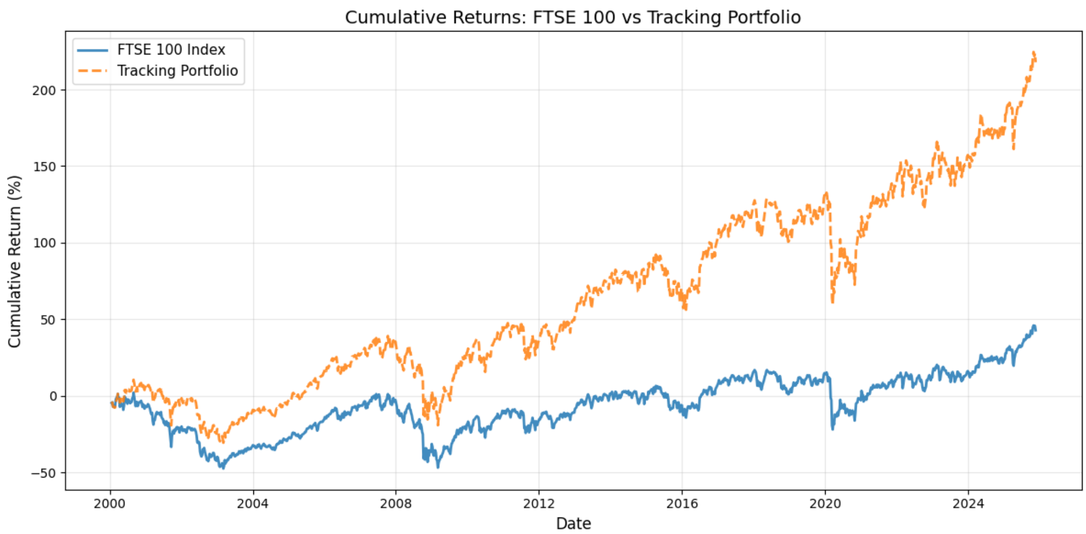


Рис. 2: Cumulative returns on FTSE 100 Index and the tracking portfolio constructed via Frank-Wolfe algorithm. The plot demonstrates close replication of index performance over time.

Пункт 4

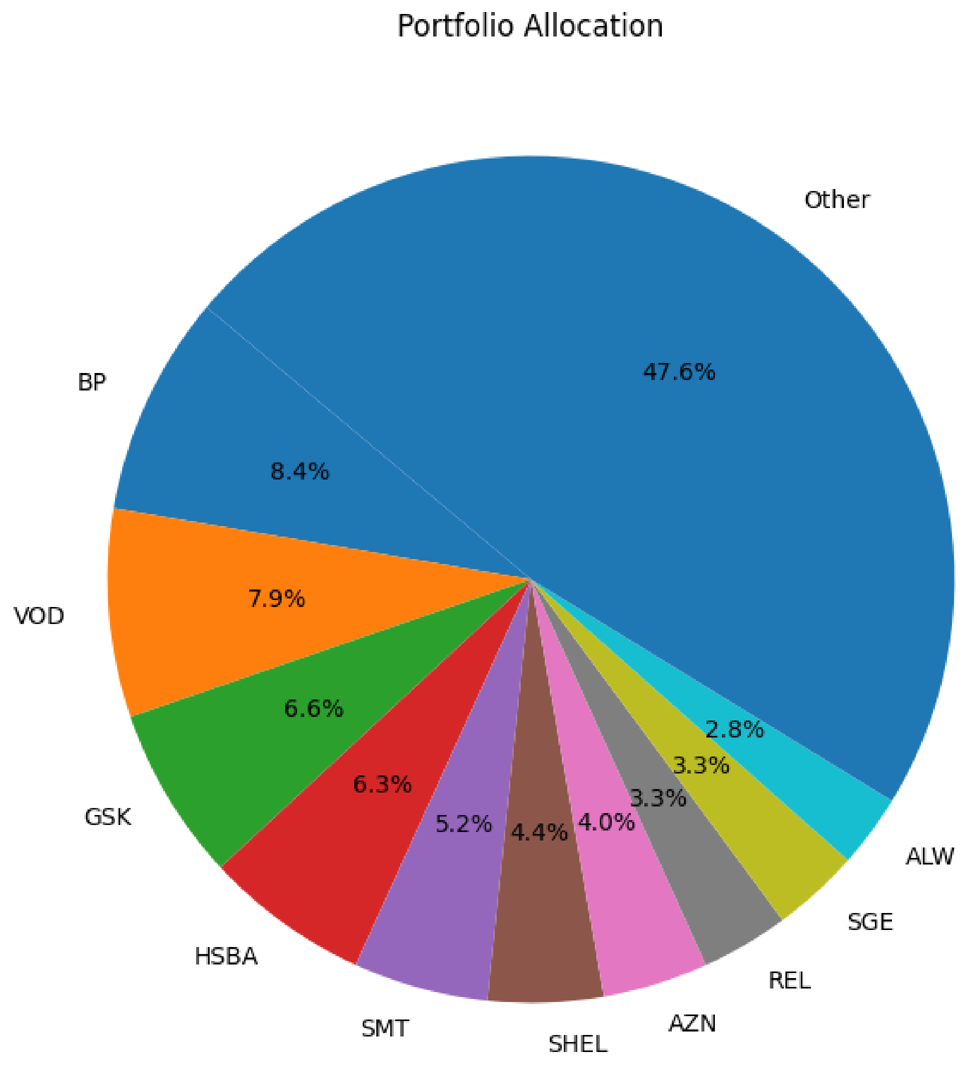


Рис. 3: Circle diagram of my portfolio turnover

Список литературы

[1] Задание №14 по курсу МФТИ «Методы оптимизации». — https://disk.yandex.ru/d/C6HBDhUXb4TdIg.

[2] Пособие по курсу МФТИ «Методы оптимизации». — https://github.com/BRAIn-Lab-teaching/OPTIMIZATION-METHODS-COURSE/blob/ПМИ$_{2025/.pdf}$.