Recap

Let \( y_0, \ldots, y^{T-1} \) be price relatives. The return of a portfolio \( x^t \in \Sigma^N \) over time horizon \([0, T]\) is

\[
\prod_{t=0}^{T-1} y^t x^t
\]

A best constant-rebalanced portfolio is a vector \( x \in \Sigma^N \) attaining

\[
\min_{x \in \Sigma^N} \frac{1}{T} \sum_{t=0}^{T-1} -\ln(y^t x)
\]

Our goal is to prove the following theorem:

**Theorem 1.** One can compute an online strategy \( x_0, \ldots, x_{T-1} \in \Sigma^N \) such that

\[
\frac{1}{T} \sum_{t=0}^{T-1} \left( \ln(y^t x^*) - \ln(y^t x^t) \right) \leq 4\rho \sqrt{\frac{\ln(N)}{T}}
\]

for any \( x^* \in \Sigma^N \), where \( \rho \) is a bound on \( \frac{y^t_i}{y^t_j} \) for all \( i, j, t \).

**Remark:** The left-hand side of the inequality is referred as average regret. Recall that the first order condition of convexity is as follows:

**Lemma 2.** Suppose that \( f : \mathbb{R}^n \to \mathbb{R} \) is differentiable and \( \text{dom}(f) \subseteq \mathbb{R}^n \) is convex. Then,

\[
f \text{ is convex } \iff \forall x, y \in \text{dom}(f) : f(y) \geq f(x) + (\nabla f(x))^T (y - x)
\]

Note that the function \( f_t : \Sigma^N \to \mathbb{R} \) with \( f_t(x) = -\ln(x^T y^t) \) is convex. For any \( x \in \Sigma^N \), set \( \rho \) as

\[
\|\nabla f_t(x)\|_\infty \leq \max \frac{y^t_i}{y^t_j} =: \rho \ \forall i, j, t.
\]
Online Portfolio Selection Using RWMA

\textbf{Theorem 3.} (Reinterpretation of [1]) Let $f_t : \Sigma^N \to \mathbb{R}$ be convex and differentiable for $t = 0, \ldots, T - 1$. One can compute $p_0, \ldots, p_{T-1} \in \Sigma^N$ online such that $\forall p^* \in \Sigma^N$,

$$\frac{1}{T} \sum_{t=0}^{T-1} [f_t(p_t) - f_t(p^*)] \leq 4\rho \sqrt{\frac{\ln(N)}{T}}, \text{where } \rho \geq \max_{t, x \in \Sigma^N} \|\nabla f_t(x)\|_\infty.$$

\textit{Proof.} To obtain such a sequence, will will apply again the randomized weighted majority algorithm. We use the following setting: The pure portfolios $e_1, \ldots, e_N$ are the $N$ experts. At time $t$

- $p_t \in \Sigma^N$ is the distribution on experts $\{1, \ldots, N\}$ (induced by the exponential weights)
- As loss vector, we choose $\ell^t = \nabla f_t(p_t)$, where $\nabla f_t(p_t) \in [-\rho, \rho]^N$.

Recall that

$$E[\hat{L}] \leq \frac{\rho \ln(N)}{\epsilon} + (1 + \epsilon)L_+^j + (1 - \epsilon)L_-^j, \text{ where } L_+^j = \sum_{t=0, \ell^t_j \geq 0} T-1 \ell^t_j, L_-^j = \sum_{t=0, \ell^t_j < 0} T-1 \ell^t_j.$$

In our setting,

$$\frac{E[\hat{L}]}{T} \leq \frac{\rho \ln(N)}{\epsilon T} + (1 + \epsilon)\frac{L_+^j}{T} + (1 - \epsilon)\frac{L_-^j}{T} = \frac{\rho \ln(N)}{\epsilon T} + (1 + \epsilon)\frac{(L_+^j + L_-^j) - 2\epsilon L_-^j}{T} \leq \frac{\rho \ln(N)}{\epsilon T} + (1 + \epsilon)\frac{(L_+^j)}{T} + 2\epsilon \rho \quad (\star)$$

2
Here we used in (⋆) that \( \frac{L_j}{T} \geq -\rho \) and hence \(-2\epsilon \frac{L_j}{T} \leq 2\epsilon \rho \). We obtain

\[
E[\hat{L}] - L_j \leq \frac{\rho \ln(N)}{\epsilon T} + 3\epsilon \rho
\]

We use this bound on the loss of the imaginary forecaster as follows:

\[
\frac{1}{T} \sum_{t=0}^{T-1} (f_t(p_t) - f_t(p^*)) \leq \frac{1}{T} \sum_{t=0}^{T-1} ((\nabla f_t(p_t))^T (p_t - p^*))
\]

\[
= \frac{1}{T} \sum_{t=0}^{T-1} ((\nabla f_t(p_t))^T p_t - (\nabla f_t(p_t))^T p^*)
\]

\[
= \frac{E[\hat{L}]}{T} - \sum_{t=0}^{T-1} (\nabla f_t(p_t))^T p^*
\]

\[
\leq \frac{E[\hat{L}] - L_j}{T}
\]

for some \( j \). In (★★) we used the first order condition \( f(y) \geq f(x) + (\nabla f(x))^T (y - x) \) (and consequently \( f(x) - f(y) \leq (\nabla f(x))^T (x - y) \)). Note that

\[
E[\hat{L}] - L_j \leq \frac{\rho \ln(N)}{\epsilon T} + 3\epsilon \rho \leq 4\rho \epsilon
\]

if we choose \( \epsilon := \sqrt{\frac{\ln(N)}{T}} \). \( \square \)

**Remark:** A proof of the First-Order Condition can be found e.g. in the book “Convex Optimization“.

Suppose we have to solve:

\[
\min_{x \in \Sigma^N} f(x) \text{ where } f : \mathbb{R}^n \to \mathbb{R} \text{ is convex and differentiable.}
\]

Use the setting from before with \( f_t := f \forall t = 0, \ldots, T - 1 \).

**Theorem 4.** With the RWMA, one can compute an \( x^* \in \Sigma^N \) such that

\[
f(x^*) - f(x) \leq \delta \text{ for all } x \in \Sigma^N \text{ with } T = \left( \frac{4\rho}{\delta} \right)^2 \ln(N).
\]

**Proof.** Use \( p_t \) from theorem before with \( f(p_t) \) minimal. \( \square \)

---

Mean Variance Portfolio Optimization

The following method is based on the diversification principle of Harry Markowitz. Note that Markowitz received the Nobel Prize in economics (1990).
Suppose that \( N \) assets are available. \( R_i \) is return of asset \( i \). \( R = \sum_{i=1}^{N} R_i x_i \) is return of portfolio \( x \in \Sigma^N \). Using \( R = 1 + r \), \( (r \) being relative return) and \( \sum_{i=1}^{N} x_i r_i \) is relative return of portfolio.

Basic notions of probability

- If \( x \) is a random variable over a finite probability space, then expected value of \( x \), \( E[x] \) or \( \bar{x} \), is defined as \( E[x] = \sum_i p_ix_i \), where \( p_i \) is the probability of \( x \) attaining the value \( x_i \).
- Linearity of expectation: \( x, y \) are random variables, \( \alpha, \beta \in \mathbb{R} \), then \( E[\alpha x + \beta y] = \alpha E[x] + \beta E[y] \).
- Variance: \( \text{Var}(x) = E[(x - \bar{x})^2] = E[x^2] - E[x]^2 \).
- Standard deviation: \( \sigma(x) = \sqrt{\text{Var}(x)} \).

Example 5. Rolling a dice \( (x \in \{1, \ldots, 6\}) \)

- \( E[x] = 3.5 \)
- \( E[x^2] = (1/6)(1+4+9+16+25+36) \)
- \( \text{Var}[x] = 2.29 \)
- Covariance: \( \text{Cov}(x, y) = E[(x - \bar{x})(y - \bar{y})] = E[xy] - \bar{x}\bar{y} \).
- Correlation: \( \text{Corr}(x, y) = \rho(x, y) = \frac{\text{Cov}(x, y)}{\sigma(x)\sigma(y)} \). Observe that \( |\rho(x, y)| \leq 1 \).
  - uncorrelated: \( \rho(x, y) = 0 \)
  - positively correlated: \( \rho(x, y) > 0 \)
  - negatively correlated: \( \rho(x, y) < 0 \)
- Variance of sum: Let \( x_1, \ldots, x_n \) be random variables. Then
  \[
  \text{Var}\left[\sum_{i=1}^{n} x_i\right] = E\left[\sum_{i=1}^{n} (x_i - \bar{x}_i)^2\right] 
  = E\left[\sum_{i,j} (x_i - \bar{x}_i)(x_j - \bar{x}_j)\right] 
  = E\left[\sum_{i,j} x_i x_j - x_i \bar{x}_j - \bar{x}_i x_j + \bar{x}_i \bar{x}_j\right] 
  = \sum_{i,j} E[x_i x_j] - \bar{x}_i \bar{x}_j 
  = \sum_{i,j} \text{Cov}(x_i, x_j)
  \]