Dynamic regularization in online portfolio selection

Dmitry B. Rokhlin

Southern Federal University
Rostov-on-Don, Russia

International Scientific Conference "New Trends of Stochastic Analysis - 2021",
Divnomorskoye, 1-5 June, 2021

- Online convex optimization problem, online gradient descent (OGD)
- Follow the leader (FTL), regularization
- Adaptive Follow the Regularized Leader (Ada-FTRL) algorithms
- Adaptive Mirror Descent (Ada-MD) algorithms
- Exponentiated Gradient (EG)
- Online Newton Step (ONS)
- Optimistic learning
- Online portfolio selection
- SOLO-FTRL algorithm for production management with transfer prices

Online learning problem

Let $\mathcal{W} \subset \mathbb{R}^d$ be a closed convex set, and let $f_1, \dots, f_T : \mathcal{W} \mapsto \mathbb{R}$ be a sequence of convex functions

$$f_t(\alpha u + (1-\alpha)v) \le \alpha f_t(u) + (1-\alpha) f_t(v), \quad \alpha \in (0,1), \quad u, v \in \mathcal{W}.$$

This sequence is unknown for the learner. At each step $t = 1, \dots, T$ an online learning algorithm

- selects a point $w_t \in \mathcal{W}$
- suffers a loss $f_t(w_t)$
- receives a feedback $g_t \in \mathbb{R}^d$
- selects the next point $w_{t+1} \in \mathcal{W}$

E.g., if we want to predict the next value of a sequence $y_t \in \mathcal{W} = [0,1]$, then typical loss functions are $f_t(w) = |w - y_t|^p$, $p = 1, p = 2$. The feedback g_t is usually the gradient or a subgradient of f_t at w_t.

The goal is to obtain a small *regret* uniformly in $w \in \mathcal{W}$:

$$\mathrm{Regret}_T(w) = \sum_{t=1}^{T} (f_t(w_t) - f_t(w)).$$

If $\sup_{w \in \mathcal{W}} \mathrm{Regret}_T(w) = o(T)$ then

$$\frac{1}{T}\sum_{t=1}^{T} f_t(w_t) - \frac{1}{T}\sum_{t=1}^{T} f_t(w_T^*) \to 0, \quad T \to \infty,$$

where w_T^* is a fixed optimal solution, taken in hindsight for each horizon T:

$$w_T^* \in \arg\max_{w \in \mathcal{W}} \sum_{t=1}^{T} f_t(w).$$

Thus, if the regret is sublinear in T, then the average loss of the algorithm is close to the average loss of the best fixed strategy taken in hindsight.

Example: the online investment problem

- $r_t = (r_{t,1}, \ldots, r_{t,d})$: the vector of returns (asset price relatives at two consecutive dates)
- $w_t = (w_{t,1}, \ldots, w_{t,d})$: the portfolio vector (fractions of wealth invested in each asset)
- $X_t = X_0 \prod_{j=1}^{t} \langle w_t, r_t \rangle$: cumulative wealth, X_0: initial wealth
- $f_t(w) = -\ln\langle w, r_t \rangle$: loss function
- $\mathcal{W} = \Delta_d = \{w \geq 0 : \sum_{t=1}^{d} w_i = 1\}$
- Regret:

$$\mathrm{Regret}_T(w) = -\sum_{t=1}^{T} \ln\langle w_t, r_t \rangle + \sum_{t=1}^{T} \ln\langle w, r_t \rangle = \ln \frac{\prod_{j=1}^{t} \langle w, r_t \rangle}{\prod_{j=1}^{t} \langle w_t, r_t \rangle}.$$

In fact, the problem of bounding this regret was posed by [Cover 1991].

Online gradient descent (OGD)

Online gradient projection algorithm [Zinkevich 2003]:

$$w_{t+1} = \Pi_{\mathcal{W}}(w_t - \eta_t g_t), \quad t = 1, \ldots, T-1,$$

$\eta_t > 0$ is a step size, g_t is a subgradient of f at w_t:

$$f(w) - f(w_t) \geq \langle g_t, w - w_t \rangle, \quad w \in \mathcal{W},$$

$\Pi_{\mathcal{W}}$ is the Euclidean projection on $\mathcal{W}$:

$$\|\Pi_{\mathcal{W}} y - y\|_2 \leq \|x - y\|_2, \quad \|x\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}.$$

Basic inequality:

$$\eta_t(f_t(w_t) - f_t(w)) \le \eta_t \langle g_t, w_t - w \rangle$$
$$\le \frac{1}{2}\|w_t - w\|_2^2 - \frac{1}{2}\|w_{t+1} - w\|_2^2 + \frac{\eta_t^2}{2}\|g_t\|_2^2,$$

$g_t \in \partial f_t(w_t)$, $w \in \mathcal{W}$. The first inequality follows from the convexity of f_t. The second inequality:

$$\|w_{t+1} - w\|_2^2 - \|w_t - w\|_2^2 = \|\Pi_{\mathcal{W}}(w_t - \eta_t g_t) - \Pi_{\mathcal{W}} w\|_2^2 - \|w_t - w\|_2^2$$
$$\le \|w_t - w - \eta_t g_t\|_2^2 - \|w_t - w\|_2^2 = -2\eta_t \langle g_t, w_t - w \rangle + \eta_t^2 \|g_t\|_2^2.$$

Reduction to online linear optimization:

$$\mathrm{Regret}_T(w) = \sum_{t=1}^T f_t(w_t) - \sum_{t=1}^T f_t(w) \le \sum_{t=1}^T \langle g_t, w_t \rangle - \sum_{t=1}^T \langle g_t, w \rangle.$$

Theorem 1

Let $D = \max_{x,y \in \mathcal{W}} \|x - y\|_2$ *be the diameter of* $\mathcal{W}$*. For* $\eta_{t+1} \leq \eta_t$*,*

$$\mathrm{Regret}_T(w) \leq \frac{D^2}{2\eta_T} + \frac{1}{2}\sum_{t=1}^{T} \eta_t \|g_t\|_2^2. \quad (1)$$

For constant step sizes $\eta_t = \eta$*,*

$$\mathrm{Regret}_T(w) \leq \frac{\|w - w_1\|^2}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{T} \|g_t\|_2^2. \quad (2)$$

Let $D < \infty$. If $\|g_t\|_2 \leq L$, then for the "optimal" value of the constant step size

$$\eta = \frac{D}{L\sqrt{T}}$$

we have

$$\mathrm{Regret}_T(w) \leq DL\sqrt{T}.$$

Online-to-batch conversion scheme

Stochastic optimization:

$$F(w) = \mathsf{E} f(w, \xi) \to \max_{w \in \mathcal{W}}$$

F is convex, $\xi_1, \ldots, \xi_T \sim \xi$, $f_t(w_t) = f(w_t, \xi_t)$,

$$\overline{w}_T = \frac{1}{T} \sum_{t=1}^{T} w_t,$$

$$F(\overline{w}_T) = \mathsf{E} f((\overline{w}_T, \xi) \leq \frac{1}{T} \sum_{t=1}^{T} \mathsf{E} f(w_t, \xi) = \frac{1}{T} \sum_{t=1}^{T} \mathsf{E} f(w_t, \xi_t)$$

$$\leq \frac{1}{T} \sum_{t=1}^{T} \mathsf{E} f(w^*, \xi_t) + \frac{1}{T} \mathsf{E}\, \mathrm{Regret}_T(w^*) = F(w^*) + \frac{1}{T} \mathsf{E}\, \mathrm{Regret}_T(w^*).$$

Lower bound: see [Hazan 2019] or [Orabona 2020]

Theorem 2

Let $\mathcal{W}$ be a bounded closed convex set. For any algorithm there exists a sequence of vectors $g_1, \dots, g_T$ with $\|g_t\|_2 \le L$ and $u \in \mathcal{W}$ such that

$$\mathrm{Regret}_T(u) = \sum_{t=1}^{T} \langle g_t, w_t \rangle - \sum_{t=1}^{T} \langle g_t, u \rangle \ge \frac{\sqrt{2} L D \sqrt{T}}{4}.$$

A construction of g_t:

$$g_t = L \varepsilon_t z,$$

where ε_t are i.i.d. Rademacher random variables: $\mathrm{P}(\varepsilon_t = 1) = \mathrm{P}(\varepsilon_t = -1) = 1/2$,

$$z = (v - w)/\|v - w\|_2, \quad \|v - w\|_2 = D.$$

Incentives to study other online learning algorithms

- finding adaptive step sizes η_t, which do not require information on the Lipschitz constant L and the horizon T;
- using variable step sizes η_t for unbounded domains;
- improvements of the estimates in the order in T for narrower classes of convex functions: strongly convex, exp-concave;
- improvements of the constants in the estimates in their dependence on the dimension d (hidden in the Lipschitz constant and domain diameter) by using non-Euclidean norms;
- incorporating predictions of the feedbacks g_t;
- parameter-free algorithms;
- scale-invariant algorithms.

Theorem 3

If $\|g_t\|_2 \leq L$ *and*

$$\eta_t = \alpha \frac{D}{L\sqrt{t}}, \quad \alpha = \frac{\sqrt{2}}{2},$$

then $\mathrm{Regret}_T(w) \leq \sqrt{2} D L \sqrt{T}$. *Furthermore, for*

$$\eta_t = \alpha \frac{D}{\sqrt{\sum_{i=1}^{t} \|g_i\|_2^2}}$$

we have

$$\mathrm{Regret}_T(w) \leq \sqrt{2} D \sqrt{\sum_{t=1}^{T} \|g_t\|_2^2}.$$

Notation

Extended functions $f : \mathbb{R}^d \mapsto \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. Effective domain: $\mathrm{dom}\, f = \{x : f(x) < +\infty\}$. A function is proper if $\mathrm{dom}\, f \neq \emptyset$ and $f(x) > -\infty$, $x \in \mathbb{R}^d$. A function $f : \mathcal{W} \mapsto \mathbb{R}$ is identified with

$$\overline{f}(x) = f(x) + \delta(x|\mathcal{W}), \quad \delta(x|\mathcal{W}) = \begin{cases} f(x), & x \in \mathcal{W}, \\ +\infty, & x \notin \mathcal{W}. \end{cases}$$

A subdifferential is the set of all subgradients u:

$$\partial f(x) = \{u : f(y) \geq f(x) + \langle y - x, u \rangle, y \in \mathbb{R}^d\}.$$

Directional derivative:

$$f'(x;z) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha z) - f(x)}{\alpha}.$$

A function f is called directionally differentiable if for any $x \in \operatorname{dom} f$ with $f(x) > -\infty$, $f'(x;z)$ exists in $[-\infty, +\infty]$ for all $z \in \mathbb{R}^d$.
Bregman divergence (for $y \in \operatorname{dom} f$):

$$B_f(x,y) = \begin{cases} f(x) - f(y) - f'(y; x-y), & \text{if } f(x) \text{ is finite,} \\ +\infty, & \text{otherwise.} \end{cases}$$

A proper convex function is directionally differentiable, and its Bregman divergence is non-negative.
Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. A convex function is called μ-strongly convex w.r.t. $\|\cdot\|$ if

$$B_f(x,y) \geq \frac{\mu}{2}\|x-y\|^2, \quad x, y \in \operatorname{dom} f.$$

Theorem 4

If the functions f_t are μ_t-strongly convex then for

$$\eta_t = \frac{1}{\sum_{j=1}^{t} \mu_j}$$

we have

$$\mathrm{Regret}_T(w) \leq \frac{1}{2} \sum_{t=1}^{T} \eta_t \|g_t\|_2^2.$$

If $\mu_t = \mu$ and the functions f_t are L-Lipschitz then

$$\mathrm{Regret}_T(w) \leq \frac{L^2}{2\mu} \sum_{t=1}^{T} \frac{1}{t} \leq \frac{L^2}{2\mu}(1 + \ln T).$$

Follow The Leader (FTL) algorithm

Notation: $c_{i:j} := \sum_{t=i}^{j} c_t$,

$$w_{t+1} = \arg\min_{w \in \mathcal{X}} f_{1:t}(w).$$

- Let f_t be 1-strongly convex w.r.t. some norm $\|\cdot\|$. Then

$$\text{Regret}_T(w) \leq \frac{1}{2}\sum_{t=1}^{T}\frac{1}{t}\|g_t\|_*^2 \leq \frac{L^2}{2}(1+\ln T).$$

  The last inequality holds true if $\|g_t\|_* \leq L$. Here $\|\cdot\|_*$ is the dual norm:

$$\|g\|_* = \sup\{\langle g, w\rangle : \|w\| \leq 1\}.$$

Example [Orabona 2020]

We want to predict a sequence $y_t \in \mathcal{X} = [0,1]$. The loss function: $f_t(w) = (w - y_t)^2/2$ is 1-strongly convex on $[0,1]$, $|g_t| = |f'(w)| = |w - y_t| \leq 1$.
FTL:

$$w_{t+1} = \arg \min_{w \in [0,1]} \sum_{j=1}^{t} \frac{1}{2}(w - y_j)^2 = \frac{1}{t} \sum_{j=1}^{t} y_j .$$

$$\mathrm{Regret}_T(w) = \sum_{t=1}^{T} \frac{1}{2}(w_t - y_t)^2 - \sum_{t=1}^{T} \frac{1}{2}(w - y_t)^2 \leq \frac{1}{2}(1 + \ln T).$$

Example [Hazan 2019]

- FTL mail fail without the strong convexity assumption.

$f_t(x) = y_t x,\ t \geq 1;\ \mathcal{X} = [-1, 1];\ y_1 = 1/2,$

$$y_t = \begin{cases} -1, & t \text{ is even} \\ 1, & t \text{ is odd} \end{cases}, \quad t \geq 2.$$

FTL: w_1 is arbitrary,

$$w_{t+1} = \arg\min_{w \in [0,1]} y_{1:t} w = \begin{cases} 1, & t \text{ is even} \\ -1, & t \text{ is odd} \end{cases} = -y_t, \quad t \geq 2.$$

Regret with respect to the fixed solution $w = 0$:

$$\text{Regret}_T(0) = w_1 y_1 + \sum_{t=2}^{T} f_t(w_t) = w_1 y_1 - \sum_{t=2}^{T} y_{t-1} y_t = w_1 y_1 + T - 1.$$

Intuition: FTL fails because it is unstable.

Example: Vovk-Azoury-Warmuth (VAW) forecaster

Online linear regression: $\mathcal{X} = \mathbb{R}^d$,

$$f_t(w) = \frac{1}{2}(y_t - \langle w, x_t \rangle)^2, \quad (x_t, y_t) \in \mathbb{R}^d \times \mathbb{R}.$$

It is assume that x_t is known before w_t is selected. Vovk-Azoury-Warmuth forecaster [Vovk 2001; Azoury and Warmuth 2001]:

$$\begin{aligned} w_{t+1} &= \arg\min_{w \in \mathbb{R}^d} \left(\frac{1}{2} \sum_{j=1}^{t} (\langle w, x_t \rangle - y_t)^2 + \frac{1}{2} \langle w, x_{t+1} \rangle^2 + \frac{\lambda}{2} \|w\|_2^2 \right) \\ &= \left(\lambda I_d + \sum_{j=1}^{t+1} x_j x_j^T \right)^{-1} \sum_{j=1}^{t} y_j x_j . \end{aligned}$$

Two regularization terms are included: one corresponds to a fictive label $y_{t+1} = 0$ and another corresponds to the Ridge regression. If $\|x_t\|_2 \leq R$, $|y_t| \leq Y$, then

$$\mathrm{Regret}_T(w) \leq \frac{\lambda}{2} \|w\|_2^2 + \frac{dY^2}{2} \ln \left(1 + \frac{R^2 T}{\lambda d} \right) .$$

Adaptive Follow the Regularized Leader (Ada-FTRL) algorithms: [Joulani, György, and Szepesvári 2020]

Two sequences of convex regularizers

$$p_1, \ldots, p_T, q_0, \ldots, q_T : \mathbb{R}^d \mapsto \mathbb{R} \cup \{+\infty\},$$

p_t, q_t which can depend on $w_1, \ldots, w_t; g_1, \ldots, g_t$. Regularizers p_t are proximal:

$$p_t(w_t) = \inf_{w \in \mathcal{W}} p_t(w).$$

Ada-FTRL:

$$w_{t+1} \in \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t}, w\rangle + p_{1:t}(w) + q_{0:t}(w)\}, \quad t = 0, \ldots, T.$$

Put also $r_t = p_t + q_{t-1}$, $t = 1, \ldots, T$.

Adaptive Mirror Descent (Ada-MD) algorithms: [Joulani, György, and Szepesvári 2020]

Two sequences of convex regularizers

$$r_1, \ldots, r_T, q_0, \ldots, q_T : \mathbb{R}^d \mapsto \mathbb{R} \cup \{+\infty\},$$

r_t, q_t which can depend on $w_1, \ldots, w_t; g_1, \ldots, g_t$.

$$\operatorname{dom} r_t \subset \operatorname{dom} r_{t-1}, \quad w_1 \in \arg\min_{w \in \mathcal{W}} q_0(w).$$

Ada-MD: $w_1 \in \arg\min_{w \in \mathcal{W}} q_0(w)$,

$$w_{t+1} \in \arg\min_{w \in \mathcal{W}} \{\langle g_t, w \rangle + q_t(w) + B_{r_{1:t}}(w, w_t)\}.$$

Put also $p_t = r_t - q_{t-1}$, $t = 1, \ldots, T$.

Mirror descent: terminology explanation from [Orabona 2020]

Let ψ be a μ-strongly convex function on $\mathcal{W}$, and let $\psi^*_{\mathcal{W}}$ be the Fenchel conjugate of $\psi_{\mathcal{W}} = \psi + \delta_{\mathcal{W}}$:

$$\psi^*_{\mathcal{W}}(v) = \sup_{u \in \mathbb{R}^d} \{\langle u, v \rangle - \psi_{\mathcal{W}}(u)\} = \sup_{u \in \mathcal{W}} \{\langle u, v \rangle - \psi(u)\}.$$

Then under some technical conditions

$$\arg\min_{w \in \mathcal{W}} \{\langle g_t, w \rangle + B_\psi(w, w_t)\} = \nabla \psi^*_{\mathcal{W}} \left(\nabla \psi_{\mathcal{W}}(w_t) - g_t\right).$$

In the non-Euclidian case the iterates w_t and gradients g_t live in different spaces. $\nabla \psi$ and $\nabla \psi^*$ are called duality mappings. OMD

- maps w_t to the dual space by $\nabla \psi$,
- performs a subgradient descent step in the dual space,
- maps the vector back to the primal space by $\nabla \psi^*$.

General regret estimates

$$\text{Ada-FTRL}: \quad \text{Regret}_T(w) \le -\sum_{t=1}^{T} B_{f_t}(w, w_t) + \sum_{t=0}^{T}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T}(p_t(w) - p_t(w_t)) - \sum_{t=1}^{T} B_{r_{1:t}}(w_{t+1}, w_t) + \sum_{t=1}^{T}\langle g_t, w_t - w_{t+1}\rangle + \sum_{t=1}^{T}\delta_t .$$

$$\text{Ada-MD}: \quad \text{Regret}_T(w) \le -\sum_{t=1}^{T} B_{f_t}(w, w_t) + \sum_{t=0}^{T}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T} B_{p_t}(w, w_t) - \sum_{t=1}^{T} B_{r_{1:t}}(w_{t+1}, w_t) + \sum_{t=1}^{T}\langle g_t, w_t - w_{t+1}\rangle + \sum_{t=1}^{T}\delta_t .$$

Assumptions

- f_t are convex,
- $g_t \in \partial f_t(w_t)$,
- $r_{1:t}$ are 1-strongly convex w.r.t some norm $\|\cdot\|_{(t)}$.

Then $\delta_t = -f'(w_t, w - w_t) + \langle g_t, w - w_t \rangle \le 0$, and

$$\begin{aligned}\langle g_t, w_t - w_{t+1} \rangle &\le \|w_t - w_{t+1}\|_{(t)} \|g_t\|_{(t),*} \\ &\le \frac{1}{2}\|w_t - w_{t+1}\|_{(t)}^2 + \frac{1}{2}\|g_t\|_{(t),*}^2 \\ &\le B_{r_{1:t}}(w_{t+1}, w_t) + \frac{1}{2}\|g_t\|_{(t),*}^2,\end{aligned}$$

where $\|\cdot\|_{(t),*}$ is the dual norm:

$$\|g\|_{(t),*} = \sup\{\langle g, w \rangle : \|w\|_{(t)} \le 1\}.$$

Theorem 5

Regret estimate for Ada-FTRL:

$$\mathrm{Regret}_T(w) \leq \sum_{t=0}^{T} (q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T} (p_t(w) - p_t(w_t)) + \sum_{t=1}^{T} \frac{1}{2} \|g_t\|_{(t),*}^2 .$$

Regret estimate for Ada-MD:

$$\mathrm{Regret}_T(w) \leq \sum_{t=0}^{T} (q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T} B_{p_t}(w, w_t) + \sum_{t=1}^{T} \frac{1}{2} \|g_t\|_{(t),*}^2 .$$

OGD as Ada-MD

Put $q_t = 0$,

$$r_1(w) = \frac{1}{2\eta_1}\|w\|_2^2, \quad r_t(w) = \left(\frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}}\right)\|w\|_2^2, \quad \eta_t \leq \eta_{t-1}$$

The function

$$r_{1:t} = \sum_{j=1}^{t} r_j(w) = \frac{1}{2\eta_t}\|w\|_2^2.$$

is 1-strongly convex w.r.t.

$$\|\cdot\|_{(t)} := \frac{1}{\sqrt{\eta_t}}\|\cdot\|_2.$$

Ada-MD gives OGD: $w_1 \in \arg\min_{w \in \mathcal{W}} q_0(w)$ is arbitrary,

$$\begin{aligned} w_{t+1} &\in \arg\min_{w \in \mathcal{W}} \{\langle g_t, w\rangle + q_t(w) + B_{r_{1:t}}(w, w_t)\} \\ &= \arg\min_{w \in \mathcal{W}} \left\{\langle g_t, w\rangle + \frac{1}{2\eta_t}\|w - w_t\|_2^2\right\} \\ &= \arg\min_{w \in \mathcal{W}} \left\{\|w\|_2^2 - 2\langle w_t - \eta_t g_t, w\rangle\right\} \\ &= \arg\min_{w \in \mathcal{W}} \|w - (w_t - \eta_t g_t)\|_2^2 = \Pi_{\mathcal{W}}(w_t - \eta_t g_t), \end{aligned}$$

$$\begin{aligned}\text{Regret}_T(w) &\le \sum_{t=0}^{T}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T} B_{p_t}(w, w_t) \\ &\quad + \sum_{t=1}^{T} \frac{1}{2}\|g_t\|_{(t),*}^2 = \sum_{t=1}^{T} B_{r_t}(w, w_t) + \sum_{t=1}^{T} \frac{\eta_t}{2}\|g_t\|_2^2.\end{aligned}$$

Since $q_t = 0$, $p_t = r_t$ and $\|\cdot\|_{(t),*} = \sqrt{\eta_t}\|\cdot\|_2$. For constant step size $\eta_t = \eta$,

$$\begin{aligned}\text{Regret}_T(w) &\le B_{r_1}(w, w_1) + \frac{\eta}{2}\sum_{t=1}^{T}\|g_t\|_2^2 \\ &= \frac{1}{2\eta}\|w - w_1\|_2^2 + \frac{\eta}{2}\sum_{t=1}^{T}\|g_t\|_2^2.\end{aligned}$$

For variable step size (formally, $\eta_0 = \infty$),

$$\sum_{t=1}^{T} B_{r_t}(w, w_t) = \sum_{t=1}^{T} \left(\frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}} \right) \|w - w_t\|_2^2$$
$$\leq D^2 \sum_{t=1}^{T} \left(\frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}} \right) = \frac{D^2}{2\eta_T}.$$

Thus,

$$\mathrm{Regret}_T(w) \leq \frac{D^2}{2\eta_T} + \frac{1}{2} \sum_{t=1}^{T} \eta_t \|g_t\|_2^2$$

if the diameter D of $\mathcal{W}$ is finite.

FTRL version of the OGD

Put $p_t = 0$,

$$q_0(w) = \frac{1}{2\eta_0}\|w\|_2^2, \quad q_t(w) = \left(\frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}}\right)\|w\|_2^2, \quad \eta_t \leq \eta_{t-1}.$$

The function

$$r_{1:t}(w) = p_{1:t}(w) + q_{0:t-1}(w) = \frac{1}{2\eta_{t-1}}\|w\|_2^2$$

is 1-strongly convex w.r.t.

$$\|\cdot\|_{(t)} = \frac{1}{\sqrt{\eta_{t-1}}}\|\cdot\|_2.$$

Ada-FTRL:

$$
\begin{aligned}
w_{t+1} &\in \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t}, w\rangle + p_{1:t}(w) + q_{0:t}(w)\} \\
&= \arg\min_{w \in \mathcal{W}} \left\{\langle g_{1:t}, w\rangle + \frac{1}{2\eta_t}\|w\|_2^2\right\} \\
&= \arg\min_{w \in \mathcal{W}} \|w + \eta_t g_{1:t}\|_2^2, \quad t \geq 0.
\end{aligned}
$$

Hence, $w_1 = \arg\min_{w \in \mathcal{W}} \|w\|_2^2$,

$$
w_{t+1} = \Pi_{\mathcal{W}}(-\eta_t g_{1:t}).
$$

A variable step size works for an unbounded domain:

$$\begin{aligned}\mathrm{Regret}_T(w) &\leq \sum_{t=0}^{T}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T}(p_t(w) - p_t(w_t)) \\ &+ \sum_{t=1}^{T} \frac{1}{2}\|g_t\|_{(t),*}^2 \leq \sum_{t=0}^{T} q_t(w) + \sum_{t=1}^{T} \frac{\eta_{t-1}}{2}\|g_t\|_2^2 \\ &= \frac{1}{2\eta_T}\|w\|_2^2 + \sum_{t=1}^{T} \frac{\eta_{t-1}}{2}\|g_t\|_2^2.\end{aligned}$$

Entropic regularizer

$\mathcal{W} = \Delta = \{w \geq 0 : \sum_{i=1}^{d} w_i = 1\}$. Negative entropy

$$\psi(w) = \sum_{i=1}^{d} w_i \ln w_i$$

is convex on Δ. Put $I(y) = \{i : y_i = 0\}$. Then

$$\begin{aligned} B_\psi(x, y) &= \psi(x) - \psi(y) - \psi'(y; x - y) \\ &= \begin{cases} \sum_{i \notin I(y)} x_i \ln \dfrac{x_i}{y_i}, & \text{if } x_i = 0, i \in I(y); \\ +\infty, & \text{if } x_i > 0 \text{ for some } i \in I(y). \end{cases} \end{aligned}$$

ψ is 1-strongly convex w.r.t. $\|x\|_1 = \sum_{i=1}^{d} |x_i|$:

$$B_\psi(x, y) \geq \frac{1}{2}\|x - y\|_1^2.$$

Exponentiated gradient (EG) [Kivinen and Warmuth 1997] as Ada-MD

Let $\mathcal{W} = \Delta$. General Ada-MD:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_t, w \rangle + q_t(w) + B_{r_{1:t}}(w, w_t)\}.$$

Put $q_t = 0$,

$$r_1(w) = \frac{\psi(w)}{\eta} = \frac{1}{\eta} \sum_{i=1}^{d} w_i \ln w_i; \quad r_t = 0, \quad t \geq 2.$$

The function $r_{1:t} = r_1$ is 1-strongly convex w.r.t.

$$\| \cdot \|_{(t)} = \frac{1}{\sqrt{\eta}} \| \cdot \|_1.$$

With the convention $\ln\frac{0}{0} = 0$ we have

$$\begin{aligned} w_{t+1} &= \arg\min_{w\in\Delta}\left\{\langle g_t, w\rangle + \frac{1}{\eta}\sum_{i=1}^{d} w_i \ln\frac{w_i}{w_{t,i}}\right\} \\ &= \arg\min_{w\in\Delta}\left\{\langle \eta g_t - \ln w_t, w\rangle + \sum_{i=1}^{d} w_i \ln w_i\right\} \\ &= \left(\frac{w_{t,i}\exp(-\eta g_{t,i})}{\sum_{j=1}^{d} w_{t,j}\exp(-\eta g_{t,j})}\right)_{i=1}^{d}, \\ \text{Regret}_T(w) &\le B_{r_1}(w, w_1) + \sum_{t=1}^{T}\frac{1}{2}\|g_t\|_{(t),*}^2 \\ &= \frac{B_\psi(w, w_1)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\|g_t\|_\infty^2, \end{aligned}$$

where the norm $\|v\|_\infty = \max_{i=1,\dots,d}|v_i|$ is dual to $\|u\|_1$.

Let w_1 be the minimum point $(1/d, \dots, 1/d)$ of ψ. Then

$$B_\psi(w, w_1) = \psi(w) - \psi(w_1) = \sum_{i=1}^{d} w_i \ln w_i + \ln d \leq \ln d.$$

If $\|g_t\|_\infty \leq L_\infty$, we get

$$\mathrm{Regret}_T(w) \leq \frac{\ln d}{\eta} + \frac{\eta}{2} L_\infty^2 T.$$

For the "optimal" $\eta = \frac{1}{L_\infty}\sqrt{\frac{2 \ln d}{T}}$,

$$\mathrm{Regret}_T(w) \leq L_\infty \sqrt{2 \ln d}\sqrt{T}.$$

For the OGD, corresponding to the Euclidian regularizer $r_1(w) = \|w\|_2^2/2$ we obtained

$$\mathrm{Regret}_T(w) \leq \frac{1}{2\eta}\|w - w_1\|_2^2 + \frac{\eta}{2}\sum_{t=1}^{T}\|g_t\|_2^2.$$

Put $w_1 = 0$ and assume that $\|g_t\|_\infty \leq L_\infty$. Since $\|w\|_2^2 \leq \sum_{i=1}^{d} w_i \leq 1$, $w \in \Delta$, we get

$$\mathrm{Regret}_T(w) \leq \frac{1}{2\eta} + \frac{\eta d}{2}L_\infty^2 T.$$

For "optimal" $\eta = \frac{1}{L_\infty}\frac{1}{\sqrt{dT}}$ the dependence on d is worse than for the enropic regularizer:

$$\mathrm{Regret}_T(w) \leq L_\infty\sqrt{d}\sqrt{T}.$$

An attempt to use a variable step size with the entropic regularizer:

$$r_1(w) = \frac{\psi(w)}{\eta_1}, \quad r_t(w) = \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\psi(w), \quad \eta_t \le \eta_{t-1}$$

leads to the estimate (where formally $\eta_0 := \infty$)

$$\mathrm{Regret}_T(w) \le \sum_{t=1}^{T}\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)B_\psi(w, w_t) + \sum_{t=1}^{T}\frac{\eta_t}{2}\|g_t\|_\infty^2.$$

In general, the first term cannot be estimated from above, since:

$$\sup_{v\in\Delta} B_\psi(w, v) = \sup_{v\in\Delta}\sum_{i=1}^{d} w_i \ln\frac{w_i}{v_i} = +\infty.$$

FTRL version of the EG

$\mathcal{W} = \Delta$. General Ada-FTRL:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t}, w\rangle + p_{1:t}(w) + q_{0:t}(w)\},$$

$r_t = p_t + q_{t-1}$, $t = 1, \dots, T$. Put $p_t = 0$,

$$q_0(w) = \frac{1}{\eta_0}(\psi(w) + \ln d), \quad q_t = \left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)(\psi(w) + \ln d), \quad \eta_t \le \eta_{t-1}.$$

Note that $q_t \ge 0$, since $0 \ge \psi(w) \ge -\ln d$, $w \in \Delta$. The functions $r_{1:t} = q_{0:t-1} = \frac{1}{\eta_{t-1}}(\psi(w) + \ln d)$ are 1-strongly w.r.t.

$$\|\cdot\|_{(t)} = \frac{1}{\sqrt{\eta_{t-1}}}\|\cdot\|_1.$$

$$\begin{aligned} w_{t+1} &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_{1:t}, w \rangle + \frac{1}{\eta_t} (\psi(w) + \ln d) \right\} \\ &= \arg\min_{w \in \mathcal{W}} \left\{ \langle \eta_t g_{1:t}, w \rangle + \sum_{i=1}^{d} w_i \ln w_i \right\} \\ &= \left(\frac{\exp(-\eta_t g_{1:t,i})}{\sum_{j=1}^{d} \exp(-\eta_t g_{1:t,j})} \right)_{i=1}^{d} . \end{aligned}$$

$$
\begin{aligned}
\mathrm{Regret}_T(w) &\le \sum_{t=0}^{T}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T}\frac{1}{2}\|g_t\|_{(t),*}^2 \\
&\le \sum_{t=0}^{T} q_t(w) + \sum_{t=1}^{T}\frac{\eta_{t-1}}{2}\|g_t\|_\infty \\
&= \frac{\psi(w) + \ln d}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_{t-1}}{2}\|g_t\|_\infty \\
&\le \frac{\ln d}{\eta_T} + \sum_{t=1}^{T}\frac{\eta_{t-1}}{2}\|g_t\|_\infty .
\end{aligned}
$$

For $\eta_t = \eta$ the same estimate was obtained for the Ada-MD version of EG. However, the Ada-FTRL version is justified for the variable step size as well.

Exp-concave functions

A function $f : \mathcal{W} \mapsto \mathbb{R}$ is called α-exp-concave, $\alpha > 0$, if the function $e^{-\alpha f}$ is concave. This is equivalent to the condition

$$\nabla^2 f(w) \succeq \alpha \nabla f(w) \nabla f(w)^T$$

for a twice differentibale function f and a domain $\mathcal{W}$ with an internal point.

- $f(w) = -\ln\langle r, w\rangle$,

$$w \in \mathcal{W} = \{w \geq 0 : \sum_{i=1}^{d} w_i = 1\}, \quad \alpha \in (0, 1].$$

- $f(w) = \langle \Sigma(Aw - b), Aw - b\rangle$, $\Sigma \succeq 0$,

$$\mathcal{W} = \left\{ w \in \mathbb{R}^d : f(w) \leq \frac{1}{\alpha} \right\}.$$

- If $f : \mathcal{W} \mapsto \mathbb{R}$ is α-exp-concave, $D = \operatorname{diam} \mathcal{W}$, $\|\nabla f(w)\|_2 \leq L$, then

$$f(u) \geq f(w) + \langle \nabla f(w), u - w \rangle + \frac{\gamma}{2} \langle \nabla f(w), u - w \rangle^2,$$
$$\gamma \leq \frac{1}{2} \min \left\{ \frac{1}{4LD}, \alpha \right\}.$$

Thus, f is strongly concave in the direction of the gradient.

Online Newton Step (ONS) [Hazan et al. 2006] as Ada-MD

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_t, w\rangle + q_t(w) + B_{r_{1:t}}(w, w_t)\}, \quad p_t = r_t - q_{t-1}.$$

Put $q_t = 0$,

$$r_1(w) = \frac{\lambda}{2}\|w\|_2^2 + \frac{1}{2}\|w\|_{A_1}^2, \quad r_t(w) = \frac{1}{2}\|w\|_{A_t}^2,$$

where

$$\|w\|_{A_t}^2 = \langle A_t w, w\rangle, \quad A_t = \gamma \nabla f_t(w_t) \nabla f_t^T(w_t).$$

We have

$$B_{r_1}(w, w_t) = \frac{\lambda}{2}\|w - w_t\|_2^2 + \frac{1}{2}\|w - w_t\|_{A_1}^2,$$

$$B_{r_j}(w, w_t) = \frac{1}{2}\|w - w_t\|_{A_j}^2, \quad j = 2, \ldots, t.$$

$r_{1:t}$ are 1-strongly convex w.r.t. the norm

$$\|w\|_{(t)} = \|w\|_{S_t} = \langle S_t w, w \rangle, \quad S_t = \lambda I_d + \sum_{j=1}^{t} A_j.$$

Ada-MD:

$$\begin{aligned} w_{t+1} &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_t, w \rangle + \frac{\lambda}{2} \|w - w_t\|_2^2 + \frac{1}{2} \sum_{j=1}^{t} \|w - w_j\|_{A_j}^2 \right\} \\ &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_t - S_t w_t, w \rangle + \frac{1}{2} \langle S_t w, w \rangle \right\} \\ &= \arg\min_{w \in \mathcal{W}} \left\| w - (w_t - S_t^{-1} g_t) \right\|_{S_t}^2 = \Pi_{\mathcal{W}}^{S_t} (w_t - S_t^{-1} g_t). \end{aligned}$$

Here $\Pi_{\mathcal{W}}^{S_t}$ is the projection on $\mathcal{W}$ with respect to the norm $\| \cdot \|_{S_t}$.

$$\begin{aligned}\mathrm{Regret}_T(w) &\leq -\sum_{t=1}^{T} B_{f_t}(w, w_t) + \sum_{t=1}^{T} B_{r_t}(w, w_t) + \frac{1}{2}\|g_t\|_{(t),*}^2 \\ &= -\sum_{t=1}^{T}\left(B_{f_t}(w, w_t) - \frac{1}{2}\|w - w_t\|_{A_t}^2\right) + \frac{\lambda}{2}\|w - w_1\|_2^2 + \sum_{t=1}^{T}\frac{1}{2}\|g_t\|_{(t),*}^2 \\ &\leq \frac{\lambda}{2}\|w - w_1\|_2^2 + \sum_{t=1}^{T}\frac{1}{2}\|g_t\|_{S_t^{-1}}^2,\end{aligned}$$

since $\|\cdot\|_{(t),*} = \|\cdot\|_{S_t^{-1}}$, and f_t are α-exp-concave:

$$\begin{aligned}B_{f_t}(w, w_t) \geq \frac{1}{2}\|w - w_t\|_{A_t}^2 &= \frac{1}{2}\langle A_t(w - w_t), w - w_t\rangle \\ &= \frac{\gamma}{2}\langle \nabla f_t(w_t)\nabla f_t^T(w_t)(w - w_t), w - w_t\rangle.\end{aligned}$$

- If $A \succeq B \succ 0$ then

$$A^{-1} \bullet (A - B) \le \ln \frac{\det A}{\det B}, \quad A \bullet B := \sum_{i,j=1}^{d} A_{ij} B_{ij} = \mathrm{Tr}\,(AB^T).$$

It follows that

$$\sum_{t=1}^{T} \|g_t\|_{S_t^{-1}}^2 = \sum_{t=1}^{T} \langle S_t^{-1} g_t, g_t \rangle = \sum_{t=1}^{T} S_t^{-1} \bullet (g_t g_t^T) = \frac{1}{\gamma} \sum_{t=1}^{T} S_t^{-1} (S_t - S_{t-1})$$

$$= \frac{1}{\gamma} \sum_{t=1}^{T} \ln \frac{\det S_t}{\det S_{t-1}} = \frac{1}{\gamma} \ln \frac{\det S_T}{\det S_0},$$

$$\mathrm{Regret}_T(w) \le \frac{\lambda}{2} \|w - w_1\|_2^2 + \frac{1}{\gamma} \ln \frac{\det S_T}{\det S_0}.$$

$$
\begin{aligned}
\langle S_T w, w\rangle &= \lambda\|w\|_2^2 + \gamma\sum_{t=1}^{T}\langle g_t g_t^T w, w\rangle \le \lambda\|w\|_2^2 + \gamma\sum_{t=1}^{T}\|g_t\|_2^2\|w\|_2^2 \\
&\le (\lambda + \gamma T L^2)\|w\|_2^2, \\
\det S_T &\le \lambda_{\max}^d(S_T) \le (\lambda + \gamma T L^2)^d, \qquad \det S_0 = \lambda^d.
\end{aligned}
$$

$$
\mathrm{Regret}_T(w) \le \frac{\lambda}{2}\|w - w_1\|_2^2 + \frac{1}{\gamma}\ln\frac{(\lambda + \gamma T L^2)^d}{\lambda^d}.
$$

Putting $\lambda = \gamma L^2$, and using the definition of γ, it can be shown that

$$
\mathrm{Regret}_T(w) \le LD\left(\frac{1}{16} + 4d\ln(1+T)\right).
$$

Online Newton Step (ONS) as Ada-FTRL

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t}, w \rangle + p_{1:t}(w) + q_{0:t}(w)\}, \quad r_t = p_t + q_{t-1}.$$

Put $q_0(w) = \frac{\lambda}{2}\|w\|_2^2$; $q_t = 0, t = 1, \ldots, T$;

$$p_t(w) = \frac{1}{2}\|w - w_t\|_{A_t}^2, \quad t = 1, \ldots, T.$$

We have

$$r_{1:t} = p_{1:t} + q_{0:t} = \frac{\lambda}{2}\|w\|_2^2 + \sum_{j=1}^{t} \frac{1}{2}\|w - w_t\|_{A_t}^2.$$

Since $\nabla^2 r_{1:t} = \lambda I_d + \sum_{j=1}^{t} A_j$, the function $r_{1:t}$ is 1-strongly convex w.r.t. $\|\cdot\|_{(t)} = \|\cdot\|_{S_t}$, $S_t = \lambda I_d + \sum_{j=1}^{t} A_j$.

Ada-FTRL (the same version as in [Orabona 2020]):

$$
\begin{aligned}
w_{t+1} &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_{1:t}, w \rangle + \sum_{j=1}^{t} \frac{1}{2} \|w - w_j\|_{A_j}^2 + \frac{\lambda}{2} \|w\|_2^2 \right\} \\
&= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_{1:t} - \sum_{j=1}^{t} A_j w_j, w \rangle + \frac{1}{2} \langle S_t w, w \rangle \right\} \\
&= \arg\min_{w \in \mathcal{W}} \left\{ \left\| S_t^{-1} g_{1:t} - S_t^{-1} \sum_{j=1}^{t} A_j w_j + w \right\|_{S_t}^2 \right\} \\
&= \Pi_{\mathcal{W}}^{S_t} \left(S_t^{-1} \sum_{j=1}^{t} A_j w_j - S_t^{-1} g_{1:t} \right).
\end{aligned}
$$

$$\begin{aligned}
\mathrm{Regret}_T(w) &\le -\sum_{t=1}^{T} B_{f_t}(w, w_t) + q_0(w) + \sum_{t=1}^{T} \frac{1}{2} p_t(w) + \frac{1}{2}\|g_t\|_{(t),*}^2 \\
&= -\sum_{t=1}^{T} \left(B_{f_t}(w, w_t) - \frac{1}{2}\|w - w_t\|_{A_t}^2 \right) + \frac{\lambda}{2}\|w\|_2^2 + \frac{1}{2}\|g_t\|_{(t),*}^2 \\
&\le \frac{\lambda}{2}\|w\|_2^2 + \frac{1}{2}\|g_t\|_{S_t^{-1}}^2 .
\end{aligned}$$

This estimate is very close to the estimate for the Ada-MD version of ONS.

Adaptive optimistic MD

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_t, w \rangle + q_t(w) + B_{r_{1:t}}(w, w_t)\}, \quad p_t = r_t - q_{t-1}.$$

Assume that we have some predictions $\widetilde{g}_t$ of g_t. Put $q_t(w) = \widetilde{q}_t(w) + \langle \widetilde{g}_{t+1} - \widetilde{g}_t, w \rangle$:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \{\langle g_t + \widetilde{g}_{t+1} - \widetilde{g}_t, w \rangle + \widetilde{q}_t(w) + B_{r_{1:t}}(w, w_t)\}.$$

$$\mathrm{Regret}_T(w) \le \sum_{t=0}^{T-1} (\widetilde{q}_t(w) - \widetilde{q}_t(w_{t+1})) + \sum_{t=1}^{T} B_{p_t}(w, w_t) + \sum_{t=1}^{T} \frac{1}{2} \|g_t - \widetilde{g}_t\|^2_{(t),*}.$$

Compare with the non-optimistic version:

$$\mathrm{Regret}_T(w) \le \sum_{t=0}^{T} (q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T} B_{p_t}(w, w_t) + \sum_{t=1}^{T} \frac{1}{2} \|g_t\|^2_{(t),*}.$$

Adaptive optimistic FTRL

General Ada-FTRL:

$$w_{t+1} \in \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t}, w\rangle + p_{1:t}(w) + q_{0:t}(w)\}, \quad r_t = p_t + q_{t-1}.$$

As for Ada-MD, put

$$q_t(w) = \widetilde{q}_t(w) + \langle \widetilde{g}_{t+1} - \widetilde{g}_t, w\rangle, \quad \widetilde{g}_0 = 0.$$

Then $q_{0:t} = \widetilde{q}_{0:t}(w) + \langle \widetilde{g}_{t+1}, w\rangle$,

$$w_{t+1} \in \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t} + \widetilde{g}_{t+1}, w\rangle + p_{1:t}(w) + \widetilde{q}_{0:t}(w)\}.$$

$$\mathrm{Regret}_T(w) \le \sum_{t=0}^{T-1}(q_t(w) - q_t(w_{t+1})) + \sum_{t=1}^{T}(p_t(w) - p_t(w_t)) + \sum_{t=1}^{T} \frac{1}{2}\|g_t - \widetilde{g}_t\|_{(t),*}^2.$$

Optimistic OGD

Ada-MD version:

$$
\begin{aligned}
w_{t+1} &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_t + \widetilde{g}_{t+1} - \widetilde{g}_t, w \rangle + \frac{1}{2\eta_t} \|w - w_t\|_2^2 \right\} \\
&= \Pi_{\mathcal{W}}(w_t - \eta_t(g_t + \widetilde{g}_{t+1} - \widetilde{g}_t)),
\end{aligned}
$$

$$
\text{Regret}_T(w) \leq \sum_{t=1}^{T} \left(\frac{1}{2\eta_t} - \frac{1}{2\eta_{t-1}} \right) \|w - w_t\|_2^2 + \sum_{t=1}^{T} \frac{\eta_t}{2} \|g_t - \widetilde{g}_t\|_2^2
$$

Ada-FTRL version:

$$
\begin{aligned}
w_{t+1} &= \arg\min_{w \in \mathcal{W}} \left\{ \langle g_{1:t} + \widetilde{g}_{t+1}, w \rangle + \frac{1}{2\eta_t} \|w\|_2^2 \right\} \\
&= \Pi_{\mathcal{W}}(-\eta_t(g_{1:t} + \widetilde{g}_{t+1})),
\end{aligned}
$$

$$
\text{Regret}_T(w) \leq \frac{1}{2\eta_T} \|w\|_2^2 + \sum_{t=1}^{T} \frac{\eta_t}{2} \|g_t - \widetilde{g}_t\|_2^2.
$$

Optimistic EG

Ada-MD version:

$$w_{t+1} = \arg\min_{w \in \Delta} \left\{ \langle g_t + \widetilde{g}_{t+1} - \widetilde{g}_t, w \rangle + \frac{1}{\eta} \sum_{i=1}^{d} w_i \ln \frac{w_i}{w_{t,i}} \right\},$$

$$w_{t+1,i} = \frac{w_{t,i} \exp\left(-\eta(g_{t,i} + \widetilde{g}_{t+1,i} - \widetilde{g}_{t,i})\right)}{\sum_{j=1}^{d} w_{t,j} \exp\left(-\eta(g_{t,j} + \widetilde{g}_{t+1,j} - \widetilde{g}_{t,j})\right)},$$

$$\text{Regret}_T(w) \leq \frac{1}{\eta} \sum_{i=1}^{d} w_i \ln \frac{w_i}{w_{1,i}} + \frac{\eta}{2} \sum_{t=1}^{T} \|g_t - \widetilde{g}_t\|_\infty^2.$$

Optimistic EG

Ada-FTRL version:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \left\{ \langle g_{1:t} + \widetilde{g}_{t+1}, w \rangle + \frac{1}{\eta_t} \sum_{i=1}^{d} w_i \ln w_i \right\}$$

$$w_{t+1,i} = \frac{\exp\left(-\eta_t (g_{1:t,i} + \widetilde{g}_{t+1,i})\right)}{\sum_{j=1}^{d} \exp\left(-\eta_t (g_{1:t,j} + \widetilde{g}_{t+1,j})\right)},$$

$$\text{Regret}_T(w) \leq \frac{\ln d}{\eta_0} + \sum_{t=1}^{T} \frac{\eta_{t-1}}{2} \|g_t - \widetilde{g}_t\|_\infty^2.$$

Online investment problem: optimistic EG

$$\text{Regret}_T(w) = \ln \frac{\prod_{j=1}^{t}\langle w, r_t\rangle}{\prod_{j=1}^{t}\langle w_t, r_t\rangle} = -\sum_{t=1}^{T} \ln\langle w_t, r_t\rangle + \sum_{t=1}^{T} \ln\langle w, r_t\rangle.$$

$$g_t = \nabla f_t(w_t) = -\frac{r_t}{\langle w_t, r_t\rangle}.$$

Ada-MD version: $w_{1,i} = 1/d$,

$$w_{t+1,i} = \frac{w_{t,i}\exp\left(-\eta z_{t,i}\right)}{\sum_{j=1}^{d} w_{t,j}\exp\left(-\eta z_{t,j}\right)},$$

$$z_t = g_t + \widetilde{g}_{t+1} - \widetilde{g}_t = -\frac{r_t}{\langle w_t, r_t\rangle} - \frac{\widetilde{r}_{t+1}}{\langle w_t, \widetilde{r}_{t+1}\rangle} + \frac{\widetilde{r}_t}{\langle w_t, \widetilde{r}_t\rangle}.$$

Ada-FTRL version: $w_{1,i} = 1/d$,

$$w_{t+1,i} = \frac{\exp\left(-\eta_t z_{t,i}\right)}{\sum_{j=1}^{d} \exp\left(-\eta_t z_{t,j}\right)},$$

$$z_t = g_{1:t} + \widetilde{g}_{t+1} = -\sum_{j=1}^{t} \frac{r_j}{\langle w_j, r_j \rangle} - \frac{\widetilde{r}_{t+1}}{\langle w_t, \widetilde{r}_{t+1} \rangle}.$$

OMD and FTRL with a fixed regularizer R, which is 1-strongly convex w.r.t. a norm $\|\cdot\|$, give the regret bounds

$$\mathrm{Regret}_T(w) \leq \frac{R(w)}{\eta} + \eta \sum_{t=1}^{T} \|g_t\|_*^2,$$

where $\eta > 0$ is a fixed learning rate. If could select the optimal $\eta = \sqrt{R(w)/\sum_{t=1}^{T} \|g_t\|_*^2}$, then we would get the bound

$$\mathrm{Regret}_T(w) \leq \sqrt{R(w) \sum_{t=1}^{T} \|g_t\|_*^2}.$$

- Parameter-free algorithms need not know an upper bound for w, but the upper bound for $\|g_t\|_*$ is assumed to be known.
- Scale-free algorithms adapt need not know an upper bound for $\|g_t\|_*$, and they are scale-invariant: the predictions are the same for g_t and cg_t, $c > 0$.

We will consider applications of two such algorithms:

- Parameter-free algorithm of [Orabona and Pál 2016] for learning with expert advice (an application to online investment).
- SOLO FTRL (Scale-free Online Linear Optimization FTRL) algorithm of [Orabona and Pál 2018] which is scale-free and can be called parameter-free since it does not need to know an upper bound for $\|g_t\|_*$ (an application to finding optimal transfer prices).

Online investment

- $r_t = (r_{t,1}, \ldots, r_{t,d})$: the vector of returns (asset price relatives at two consecutive dates)
- $p_t = (p_{t,1}, \ldots, p_{t,d})$: the portfolio vector (fractions of wealth invested in each asset)
- $X_t = X_0 \prod_{j=1}^{t} \langle p_t, r_t \rangle$: cumulative wealth, X_0: initial wealth
- $f_t(p) = -\ln\langle p, r_t \rangle$: loss function
- $\{p \geq 0 : \sum_{t=1}^{d} p_i = 1\}$
- Regret:

$$\mathrm{Regret}_T(q) = -\sum_{t=1}^{T} \ln\langle p_t, r_t \rangle + \sum_{t=1}^{T} \ln\langle q, r_t \rangle = \ln \frac{\prod_{j=1}^{t} \langle q, r_t \rangle}{\prod_{j=1}^{t} \langle p_t, r_t \rangle}.$$

Standard data sets [Li et al. 2015]

Dataset	Market	Region	Time Frame	# Periods	# Assets
NYSE (O)	Stock	US	Jul. 3rd 1962 Dec. 31st 1984	5651	36
NYSE (N)	Stock	US	Jan. 1st 1985 Jun. 30th 2010	6431	23
TSE	Stock	CA	Jan. 4th 1994 Dec. 31st 1998	1259	88
SP500	Stock	US	Jan. 2nd 1998 Jan. 31st 2003	1276	25
MSCI	Index	Global	Apr. 1st 2006 Mar. 31st 2010	1043	24
DJIA	Stock	US	Jan. 14th 2001 Jan. 14th 2003	507	30

	NYSE (O)	NYSE (N)	DJIA	TSE	SP500	MSCI
BCRP	250.60	120.32	1.24	6.78	4.07	1.51
Uniform	27.07	31.55	0.81	1.60	1.65	0.93
EG	27.09	31.00	0.81	1.59	1.63	0.93
ONS	109.19	21.59	1.53	1.62	3.34	0.86
OEG (perfect)	29.81	34.82	0.82	1.65	1.68	0.93

- BCRP (Best constant rebalanced portfolio): optimal portfolio weights taken in hindsight;
- Uniform: uniform portfolio weights $w_t = (1/d, \dots, 1/d)$.
- EG (Exponential Gradient) algorithm with the learning rate $\eta = 0.05$ as suggested by [Helmbold et al. 1998];
- ONS (Online Newton Step) algorithm with parameters $\eta = 0$, $\beta = 1$, $\delta = 1/8$ as suggested by [Agarwal et al. 2006];
- OEG (perfect): MD-version of the OEG with (unrealisable) perfect predictions $\widetilde{g}_{t+1} = g_{t+1}$.

On-Line Moving Average Reversion (OLMAR): [Li et al. 2015]

Asset prices: $s_t = (s_t^1, \ldots, s_t^d)$, price relatives (returns): $r_t = s_t / s_{t-1}$ (componentwise), predictions:

- MAR-1:

$$\widetilde{r}_{t+1} = \frac{1}{k} \frac{s_t + s_{t-1} + \cdots + s_{t-k+1}}{s_t}$$

- MAR-2:

$$\begin{aligned} \widetilde{r}_{t+1} &= \frac{\alpha s_t + (1-\alpha)\alpha s_{t-1} + (1-\alpha)^2 \alpha s_{t-2} + \cdots + (1-\alpha)^{t-1} \alpha s_1}{s_t} \\ &= \alpha \mathbb{1} + (1-\alpha) \frac{\widetilde{r}_t}{r_t} \end{aligned}$$

$$p_{t+1/2} = \arg\min \|p - p_t\|_2^2, \quad \langle p, \widetilde{r}_{t+1} \rangle \geq \varepsilon, \quad \langle p, \mathbb{1} \rangle = 1,$$

$$p_{t+1} = \arg\min_{p \in \Delta_d} \min \|p - p_{t+1/2}\|_2^2, \quad \Delta_d = \{p \geq 0 : \sum_{t=1}^{d} p_i = 1\}.$$

Expert strategies

Let $\widetilde{r}_{t+1}$ be the predictions of MAR-1 or MAR-2 with a fixed parameter α or m. Consider simple expert portfolio strategies:

$$j_t = \arg\max \widetilde{r}_{t+1,j}; \qquad q_{t,j_t} = 1, \quad q_{t,i} = 0, \quad i \neq j_t.$$

That is, at each step invest all wealth in the stock with the largest predicted return. Construct m_1 such strategies of MAR-1 type and m_2 such strategies of MAR-2 type, based on the predictions

$$\widetilde{r}^{(k)}_{t+1} = \frac{1}{k}\frac{s_t + s_{t-1} + \cdots + s_{t-k+1}}{s_t}, \quad k \in \{2, \ldots, m_1 + 1\}$$

$$\widetilde{r}^{(\alpha)}_{t+1} = \alpha\mathbb{1} + (1-\alpha)\frac{\widetilde{r}_t}{r_t}, \quad \alpha \in \{k/m_2 : k = 0, \ldots, m_2\}$$

respectively.

Parameter-free algorithm for learning with expert advice

$$\text{Regret}_T(w) = \sum_{t=1}^{T} \langle g_t, w_t \rangle - \sum_{t=1}^{T} \langle g_t, w \rangle,$$

$\mathcal{W} = \Delta_m = \{w \in \mathbb{R}_+^m : \|w\|_1 = 1\}$: learning with Expert Advice, $g_t \in [0,1]^m$. Let $\pi \in \Delta_m$. The algorithm [Orabona and Pál 2016]:

$$u_{t,i} = \frac{\sum_{j=1}^{t-1} \widetilde{g}_{j,i}}{t + T/2} \left(1 + \sum_{j=1}^{t-1} \widetilde{g}_{j,i} u_{j,i} \right),$$

$$\widehat{w}_{t,i} = \pi_i u_{t,i}^+,$$

$$w_t = \begin{cases} \widehat{w}_t / \|\widehat{w}_t\|_1, & \|p_t\|_1 > 0, \\ \pi, & \|p_t\|_1 = 0, \end{cases}$$

$$\widetilde{g}_{t,i} = \begin{cases} g_{t,i} - \langle g_t, w_t \rangle, & u_{t,i} > 0, \\ (g_{t,i} - \langle g_t, w_t \rangle)^+, & u_{t,i} \leq 0. \end{cases}$$

Regret bound [Orabona and Pál 2016]:

$$\mathrm{Regret}_T(w) \le \sqrt{3T(3 + \mathrm{D}(w||\pi))},$$

where $\mathrm{D}(p||q) = \sum_{i=1}^m p_i \ln(p_i/q_i)$ is the Kullback-Leibler divergence. Consider $\lceil \varepsilon m \rceil$ best experts. Assume that they are ordered from the best to the worst. Put

$$w_\varepsilon = (1/\lceil \varepsilon m \rceil, \dots, 1/\lceil \varepsilon m \rceil, 0, \dots, 0),$$

and let i_ε be the number of the $\lceil \varepsilon m \rceil$-best expert. Then for $\pi = (1/m, \dots, 1/m)$ we get the quantile bound (see [Orabona 2020])

$$\begin{aligned}\mathrm{Regret}_T(\varepsilon) = \sum_{t=1}^T \langle g_t, w_t \rangle - \sum_{t=1}^T g_{t,i_\varepsilon} &\le \sum_{t=1}^T \langle g_t, w_t \rangle - \sum_{t=1}^T \langle g_t, w_\varepsilon \rangle \\ &\le \sqrt{3T(3 + \mathrm{D}(w_\varepsilon||\pi)} \le \sqrt{3T(3 + \ln(1/\varepsilon))},\end{aligned}$$

which does not depend on the number m of experts.

Let

$$\underline{a} \le \min_{t,i} r_{t,i}, \quad \max_{t,i} r_{t,i} \le \overline{a}.$$

Take any portfolio strategies $q_t^{(i)} \in \Delta_d$, $i = 1, \ldots, m$, and put

$$g_{t,i} = \frac{\ln \overline{a} - \ln \langle q_t^{(i)}, r_t \rangle}{\ln \overline{a} - \ln \underline{a}} \in [0, 1].$$

We have

$$\sum_{t=1}^{T} \sum_{i=1}^{m} \langle w_t, g_t \rangle - \sum_{t=1}^{T} \sum_{i=1}^{m} \langle w, g_t \rangle = \sum_{t=1}^{T} \sum_{i=1}^{m} (w_{t,i} - w_i) \frac{\ln \overline{a} - \ln \langle q_t^{(i)}, r_t \rangle}{\ln \overline{a} - \ln \underline{a}}$$

$$= -\frac{1}{\ln \overline{a} - \ln \underline{a}} \sum_{t=1}^{T} \sum_{i=1}^{m} (w_{t,i} - w_i) \ln \langle q_t^{(i)}, r_t \rangle \le \sqrt{3T(3 + \ln(1/\varepsilon))}.$$

Let $X_T = \prod_{t=1}^{T} \langle \sum_{i=1}^{m} w_{t,i} q_t^{(i)}, r_t \rangle$ be the wealth obtained by the combination of experts' portfolios according to the algorithm. Then

$$\ln X_T \geq \sum_{t=1}^{T} \sum_{i=1}^{m} w_i \ln \langle q_t^{(i)}, r_t \rangle - \sqrt{3T(3 + \mathsf{D}(w||\pi)} \ln \left(\frac{\overline{a}}{\underline{a}} \right).$$

In particular,

$$\ln \frac{X_T}{X_{T,\varepsilon}} \geq -\sqrt{3T(3 + \ln(1/\varepsilon)} \ln \left(\frac{\overline{a}}{\underline{a}} \right),$$

where $X_{T,\varepsilon}$ is the wealth of the $\lceil \varepsilon m \rceil$-best expert portfolio.

	NYSE (O)	NYSE (N)	DJIA	TSE	SP500	MSCI
BCRP	250.60	120.32	1.24	6.78	4.07	1.51
OLMAR-1	$7.2 \cdot 10^{16}$	$4.32 \cdot 10^{8}$	2.20	54.52	17.05	14.36
OLMAR-2	$3.09 \cdot 10^{18}$	$6.80 \cdot 10^{8}$	1.41	138.59	13.66	16.40
OLMAR-1-ext	$5.04 \cdot 10^{16}$	$4.06 \cdot 10^{8}$	2.49	70.94	15.61	13.63
OLMAR-2-ext	$3.64 \cdot 10^{18}$	$4.46 \cdot 10^{8}$	1.26	310.45	15.40	17.60
OLMAR-1-free	$1.03 \cdot 10^{17}$	$3.28 \cdot 10^{8}$	1.72	301.05	17.79	16.45
OLMAR-2-free	$3.14 \cdot 10^{17}$	$8.54 \cdot 10^{8}$	1.36	456.82	6.46	16.70
OLMAR-free	$2.10 \cdot 10^{17}$	$6.77 \cdot 10^{8}$	1.91	434.76	10.27	16.69

- Parameters for OLMAR [Li et al. 2015]: $k = 5$, $\alpha = 0.3$, $\varepsilon = 10$.
- OLMAR-ext(reme): one simple expert strategy with the same parameters.
- OLMAR-1-free: MAR-1 experts, $k = 2, \ldots, 6$ $(m_1 = 5)$
- OLMAR-2-free: MAR-2 experts, $\alpha \in \{0, 0.1, \ldots, 0.9, 1\}$ $(m_2 = 11)$
- OLMAR-free: $m_1 = 5$ MAR-1 experts and $m_2 = 11$ MAR-2 experts

Production management with transfer prices

Consider a firm consisting from n production and m sales divisions. There are d commodities produced by each production division. The same commodities are saled by each sales division.

- $f_i : X_i \mapsto \mathbb{R}_+$, $i = 1, \ldots, m$: revenue functions of the sales divisions,
- $g_i : Y_i \mapsto \mathbb{R}_+$, $i = 1, \ldots, n$ cost functions of the production divisions,
- $x_i \in X_i$: the amounts of commodities to be sold by i-th sales division,
- $y_i \in Y_i$: the amounts of commodities to be produced by i-th production division,
- X_i, Y_i: some convex subsets of $\mathbb{R}^d_+ = \{x \geq 0 : x \in \mathbb{R}^d\}$.

Under an unrealistic assumption that the functions f_i, g_i are completely known, the firm manager can solve the profit maximization problem

$$F(x, y) = \sum_{i=1}^{m} f_i(x_i) - \sum_{i=1}^{n} g_i(y_i) \to \max_{(x,y)\in\mathcal{W}}, \qquad (3)$$

$$\mathcal{W} = \left\{(x, y) \in Z : \sum_{i=1}^{m} x_i = \sum_{j=1}^{n} y_j\right\}, \quad Z = \prod_{i=1}^{m} X_i \times \prod_{j=1}^{n} Y_j, \qquad (4)$$

and ensure an optimal firm performance by assigning to each division the related component of an optimal solution. The constraint $(x, y) \in \mathcal{W}$ requires that the total production equals to the total sales in each commodity: an equilibrium between the supply and demand at the firm level.

Under convexity assumptions there is also a more economic way to achieve the same goal. The firm can announce the commodity *transfer price* vector $\lambda_t \in \mathbb{R}^d_+$ with the obligation to buy the commodities at these prices from the production divisions, and sell them to the sales divisions.
Optimal division (agent) reactions are defined by

$$\widetilde{x}_i(\lambda) \in \arg\max_{x_i \in X_i} (f_i(x_i) - \langle \lambda, x_i \rangle), \quad i = 1, \dots, m,$$

$$\widetilde{y}_i(\lambda) \in \arg\max_{y_i \in Y_i} (\langle \lambda, y_i \rangle - g_i(y_i)), \quad i = 1, \dots, n,$$

We will say that the plan $\widetilde{z}(\lambda) = (\widetilde{x}(\lambda), \widetilde{y}(\lambda))$ is stimulated by the transfer price vector λ.

Assumption 1. The sets X_i, Y_i are convex, compact, and

$$[0,\varepsilon]^d \subset X_i \subset [0,a]^d, \quad [0,\varepsilon]^d \subset Y_i \subset [0,a]^d$$

with some $\varepsilon > 0$, $a > 0$.

Assumption 2. The functions $f_i : X_i \mapsto \mathbb{R}_+$ (resp., $g_i : Y_i \mapsto \mathbb{R}_+$) are σ_i'-strongly concave (resp., σ_i''-strongly convex), non-decreasing in each variable, and $f_i(0) = g_i(0) = 0$.

Assumption 3. The functions f_i (resp., g_i) are K_i'-Lipschitz (resp., K_i''-Lipschitz).

Theorem 6

An admissible point $z^* = (x^*, y^*) \in \mathcal{W}$ is an optimal solution of the problem (3), (4) if and only if it is stimulated by some transfer price vector $\lambda^* \in \mathbb{R}_+^d$: $z^* = \widetilde{z}(\lambda^*)$.

The price vector λ^*, mentioned in Theorem 6, is an optimal solution of the dual minimization problem

$$G(\lambda) := \sup_{(x,y)\in X\times Y} L(x,y,\lambda) = \sum_{i=1}^{m} v_i(\lambda) - \sum_{i=1}^{n} u_i(\lambda) \to \min_{\lambda\in\mathbb{R}^d}, \qquad (5)$$

where v_i and u_i are Fenchel convex and concave conjugates of g_i and f_i respectively:

$$v_i(\lambda) = \sup_{y_i\in Y_i} (\langle\lambda, y_i\rangle - g_i(y_i)), \quad u_i(\lambda) = \inf_{x_i\in X_i} (\langle\lambda, x_i\rangle - f_i(x_i)).$$

The dual problem (5) is solvable and there is no duality gap: $F(x^*, y^*) = G(\lambda^*)$.

We have

$$\nabla G(\lambda) = \Delta\widetilde{z}(\lambda) := \sum_{i=1}^{n} \widetilde{y}_i(\lambda) - \sum_{i=1}^{m} \widetilde{x}_i(\lambda).$$

Let us consider the gradient descent algorithm with some step sizes $\eta_t > 0$:

$$\lambda_{t+1} = \lambda_t - \eta_t \nabla G(\lambda_t) = \lambda_t - \eta_t \Delta\widetilde{z}(\lambda_t). \tag{6}$$

Observing the agents' reactions on the transfer price vector λ_t, the firm manager can iteratively update λ_t according to the law of supply and demand.

It is easy to show that G is κ-smooth:

$$\|\nabla G(x) - \nabla G(y)\| \le \kappa\|x - y\|, \quad \kappa = \sum_{i=1}^{m} \frac{1}{\sigma_i'} + \sum_{i=1}^{n} \frac{1}{\sigma_i''}.$$

Nesterov's fast gradient descent algorithm [Nesterov 1983] in the form considered in [Su, Boyd, and Candès 2016]:

$$\lambda_t = \mu_{t-1} - \eta \nabla G(\mu_{t-1}) = \mu_{t-1} - \eta \Delta \widetilde{z}(\mu_{t-1}), \tag{7}$$

$$\mu_t = \lambda_t + \frac{t-1}{t+2}(\lambda_t - \lambda_{t-1}), \quad \mu_0 = \lambda_0 \in \mathbb{R}^d. \tag{8}$$

Note that this algorithm comes from a slightly general family than (6). If

$$\eta \leq 1/\kappa = \sum_{i=1}^{m} \frac{1}{\sigma_i'} + \sum_{i=1}^{n} \frac{1}{\sigma_i''}$$

then (see [Su, Boyd, and Candès 2016])

$$G(\lambda_t) - G(\lambda^*) \leq \frac{2\|\lambda_0 - \lambda^*\|^2}{\eta(t+1)^2}.$$

Theorem 7

If $\eta \in (0, 1/\kappa]$, then for the transfer price sequence λ_t, generated by the fast gradient descent (7), (8), we have

$$|F(\tilde{z}(\lambda_t)) - F(z^*)| \le \frac{2K}{\sqrt{\sigma\eta}} \frac{\|\lambda_0 - \lambda^*\|}{t+1},$$

$$\|\Delta\tilde{z}(\lambda_t)\| \le 2\sqrt{\frac{\kappa}{\eta}} \frac{\|\lambda_0 - \lambda^*\|}{t+1},$$

where

$$\sigma = \min\left\{\min_{1\le i\le m} \sigma_i', \min_{1\le i\le n} \sigma_i''\right\},$$

$$K = \left(\sum_{i=1}^{m}(K_i')^2 + \sum_{i=1}^{n}(K_i'')^2\right)^{1/2}.$$

To implement the algorithm (7), (8) one needs to know the smoothness parameter κ of G. However, in practice it is unknown.

SOLO-FTRL algorithm [Orabona and Pál 2018]

FTRL with Varying Regularizer:

$$w_t = \arg\min_{w \in \mathcal{W}} \{\langle g_{1:t-1}, w \rangle + R_t(w)\}, \quad t = 1, \ldots, T.$$

SOLO-FTRL (Scale-free Online Linear Optimization FTRL) algorithm:

$$R_t(w) = R(w) \sqrt{\sum_{j=1}^{t-1} \|g_j\|_*^2}.$$

Theorem 8 (Orabona and Pál 2018)

Let $\mathcal{W}$ be a closed convex set, $D = \sup_{u,v \in \mathcal{W}} \|u - v\| \in [0, \infty]$. Suppose that $R : \mathcal{W} \mapsto \mathbb{R}$ is a non-negative μ-strongly convex function w.r.t. $\|\cdot\|$. Then the regret of the SOLO-FTRL algorithm satisfies

$$\mathrm{Regret}_T(w) \leq \left(R(w) + \frac{2.75}{\mu}\right)\sqrt{\sum_{j=1}^{T} \|g_j\|_*^2} + 3.5 \min\left\{\frac{\sqrt{T-1}}{\mu}, D\right\} \max_{1 \leq t \leq T} \|g_t\|_*.$$

The SOLO FTRL with $R(w) = \|w\|_2^2/2$ does not require any parameter knowledge:

$$\lambda_0 = 0, \quad L_0 = 0,$$
$$\lambda_t = \arg\min_{\lambda \in \mathbb{R}^d} \left(\langle L_{t-1}, \lambda \rangle + \sqrt{\sum_{j=1}^{t-1} \|\nabla G(\lambda_j)\|^2} \cdot \|\lambda\|^2/2 \right), \quad t \geq 1, \qquad (9)$$
$$L_t = L_{t-1} + \nabla G(\lambda_t).$$

Solving the optimization problem (9), we get

$$\lambda_t = -\frac{L_{t-1}}{\sqrt{\sum_{j=1}^{t-1} \|\nabla G(\lambda_j)\|^2}} = -\frac{\sum_{j=1}^{t-1} \Delta\widetilde{z}(\lambda_j)}{\sqrt{\sum_{j=1}^{t-1} \|\Delta\widetilde{z}(\lambda_j)\|^2}}, \quad \lambda_0 = 0. \qquad (10)$$

$$\mathrm{Regret}_T(\lambda) \leq \left(\|\lambda\|^2/2 + 2.75\right)\sqrt{\sum_{t=1}^{T}\|\nabla G(\lambda_t)\|^2} + 3.5\sqrt{T-1}\max_{t\leq T}\|\nabla G(\lambda_t)\|.$$

Let us estimate $\nabla G(\lambda)$:

$$\|\nabla G(\lambda)\| \leq \sum_{i=1}^{n}\|\nabla v_i(\lambda)\| + \sum_{i=1}^{m}\|\nabla u_i(\lambda)\| = \sum_{i=1}^{n}\|\widetilde{y}_i(\lambda)\| + \sum_{i=1}^{m}\|\widetilde{x}_i(\lambda)\|\|\| \leq (m+n)a\sqrt{d}. \tag{11}$$

For $\overline{\lambda}_T = \frac{1}{T}\sum_{t=1}^{T}\lambda_t$ we get

$$G(\overline{\lambda}_T) - G(\lambda^*) \leq \frac{1}{T}\sum_{t=1}^{T}(G(\lambda_t) - G(\lambda^*)) = \frac{\mathrm{Regret}_T(\lambda^*)}{T} \leq \frac{(m+n)a\sqrt{d}}{\sqrt{T}}\left(\frac{\|\lambda^*\|^2}{2} + 6.25\right). \tag{12}$$

Theorem 9

For the average transfer price vector

$$\overline{\lambda}_T = \frac{1}{T}\sum_{t=1}^{T}\lambda_t,$$

generated by the SOLO FTRL algorithm (10), we have

$$F(z^*) - F(\widetilde{z}(\overline{\lambda}_T)) \le K\sqrt{\frac{(m+n)a}{\sigma}}\sqrt{\|\lambda^*\|^2 + 12.5}\frac{d^{1/4}}{T^{1/4}},$$

$$\|\Delta\widetilde{z}(\overline{\lambda}_T)\| \le \sqrt{\kappa(m+n)a}\sqrt{\|\lambda^*\|^2 + 12.5}\frac{d^{1/4}}{T^{1/4}}.$$

The iterations are uniformly bounded:

$$-1 \le \lambda_{t,k} \le K' + 1, \quad K' = \max_{i=1,\dots,m} K_i'.$$

It should be mentioned that the convergence rate of this algorithm is much more slow than the convergence rate of the fast gradient descent.

Dynamic problem

Consider a sequence of time dependent revenue and cost functions

$$f_{t,i} : X_i \mapsto \mathbb{R}_+, \quad g_{t,i} : Y_i \mapsto \mathbb{R}_+,$$

and the sequence of profit maximization problems

$$F_t(x,y) = \sum_{i=1}^{m} f_{t,i}(x_i) - \sum_{i=1}^{n} g_{t,i}(y_i) \to \max_{(x,y)\in\mathcal{W}}, \tag{13}$$

$$\mathcal{W} = \left\{(x,y) \in Z : \sum_{i=1}^{m} x_i = \sum_{j=1}^{n} y_j\right\}, \quad Z = \prod_{i=1}^{m} X_i \times \prod_{j=1}^{n} Y_j, \tag{14}$$

where $\mathcal{W}$ is defined by (14). We assume that the Assumptions $1-3$ are still satisfied for $f_{t,i}$, $g_{t,i}$ instead of f_i, g_i, wherein all constants ε, a, σ_i', σ_i'', K_i', K_i'' are independent of t.

Recall that the notation $X_t = O_{\mathsf{P}}(c_t)$, where X_t are random variables, P is a probability measure, and $c_t > 0$ are constants, means that the sequence X_t/c_t is stochastically bounded. That is, for any $\varepsilon > 0$ there exists $C > 0$, $t_0 > 0$ such that

$$\mathsf{P}(|X_t|/c_t \geq C) \leq \varepsilon, \quad t \geq t_0.$$

Assume additionally that the revenue and cost functions are of the form

$$f_{t,i}(x_i) = f_i(x_i, \xi_{t,i}), \quad g_{t,i}(y_i) = g_i(y_i, \eta_{t,i}), \tag{15}$$

where $(\xi_t, \eta_t) \in \Theta \subset \mathbb{R}^m \times \mathbb{R}^n$ is a sequence of i.i.d. random vectors such that

$$\mathsf{E} \sup_{x_i \in X_i} f_i(x_i, \xi_{t,i}) < \infty, \quad \mathsf{E} \sup_{y_i \in Y_i} g_i(y_i, \eta_{t,i}) < \infty. \tag{16}$$

Theorem 10

Under the assumptions (15), (16) the price sequence generated by the SOLO FTRL algorithm (10) ensures no-regret learning with respect to the best possible plan sequence z_t^*:

$$\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t)) \to 0 \quad \text{a.s,} \quad T \to \infty,$$

and the estimate

$$\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t)) = O_{\mathsf{P}}\left(\frac{1}{T^{1/4}}\right).$$

The equilibrium condition is satisfied on average:

$$\frac{1}{T}\sum_{t=1}^{T}\Delta\widetilde{z}_t(\lambda_t) = O_{\mathsf{P}}\left(\frac{1}{T^{1/4}}\right).$$

Proof. There exists $\overline{\lambda}$, satisfying the equation

$$\mathsf{E}\nabla G_t(\overline{\lambda}) = 0.$$

From the definition of the dual objective functions we get

$$G_t(\lambda) = F_t(\widetilde{z}_t(\lambda)) - \langle \lambda, \Delta\widetilde{z}_t(\lambda)\rangle, \quad \Delta\widetilde{z}_t(\lambda) := \sum_{i=1}^{m} \widetilde{y}_{t,i}(\lambda) - \sum_{i=1}^{n} \widetilde{x}_{t,i}(\lambda).$$

Since $\Delta\widetilde{z}_t(\lambda) = \nabla G_t(\lambda)$, and λ_t^* is a minimum point of G_t, using the strong duality we obtain the inequality

$$\begin{aligned}
&F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t)) = G_t(\lambda_t^*) - G_t(\lambda_t) - \langle \lambda_t, \nabla G_t(\lambda_t)\rangle \\
&\le -\langle \lambda_t, \nabla G_t(\lambda_t)\rangle = -\langle \lambda_t, \nabla G_t(\lambda)\rangle + \langle \lambda_t, \nabla G_t(\lambda) - \nabla G_t(\lambda_t)\rangle \\
&\le -\langle \lambda_t, \nabla G_t(\lambda)\rangle + \|\lambda_t\| \cdot \|\nabla G_t(\lambda) - \nabla G_t(\lambda_t)\|
\end{aligned}$$

for any λ.

$$
\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t))) \\
&\leq -\frac{1}{T}\sum_{t=1}^{T}\langle \lambda_t, \nabla G_t(\overline{\lambda})\rangle + \frac{1}{T}\sum_{t=1}^{T}\|\lambda_t\| \cdot \|\nabla G_t(\overline{\lambda}) - \nabla G_t(\lambda_t)\| \\
&\leq -\frac{1}{T}\sum_{t=1}^{T}\langle \lambda_t, \nabla G_t(\overline{\lambda})\rangle + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\|\lambda_t\|^2}\sqrt{\frac{1}{T}\sum_{t=1}^{T}\|\nabla G_t(\overline{\lambda}) - \nabla G_t(\lambda_t)\|^2}.
\end{aligned}
$$

Since the functions G_t are κ-smooth, it follows that

$$
G_t(\lambda_t) - G_t(\overline{\lambda}) \geq \langle \nabla G_t(\overline{\lambda}), \lambda_t - \overline{\lambda}\rangle + \frac{1}{2\kappa}\|\nabla G_t(\lambda_t) - \nabla G_t(\overline{\lambda})\|^2.
$$

Using the estimate for the regret:

$$\sum_{t=1}^{T}(G_t(\lambda_t) - G_t(\lambda)) \le (m+n)a\sqrt{d}\left(\frac{\|\lambda\|^2}{2} + 6.25\right)\sqrt{T},$$

we conclude that if

$$\frac{1}{T}\sum_{t=1}^{T}\|\nabla G_t(\bar{\lambda}) - \nabla G_t(\lambda_t)\|^2 \ge \frac{A^2}{4b^2 d T^{1/2}}$$

then

$$-\frac{1}{T}\sum_{t=1}^{T}\langle\lambda_t, \nabla G_t(\bar{\lambda})\rangle \ge \frac{A^2}{8\kappa b^2 d\sqrt{T}} - \frac{1}{T}\sum_{t=1}^{T}(G_t(\lambda_t) - G_t(\bar{\lambda}))$$
$$\ge \frac{A^2}{16\kappa b^2 d}\frac{1}{\sqrt{T}}$$

for sufficiently large A.

Hence,

$$\mathsf{P}\left(\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t))) \geq \frac{A}{T^{1/4}}\right) \leq \mathsf{P}\left(-\frac{1}{T}\sum_{t=1}^{T}\langle\lambda_t, \nabla G_t(\bar{\lambda})\rangle \geq \frac{A}{2T^{1/4}}\right)$$

$$+\,\mathsf{P}\left(-\frac{1}{T}\sum_{t=1}^{T}\langle\lambda_t, \nabla G_t(\bar{\lambda})\rangle \geq \frac{A^2}{16\kappa b^2 d}\frac{1}{\sqrt{T}}\right).$$

Furthermore, in view of the estimate

$$|\langle\lambda_t, \nabla G_t(\bar{\lambda})\rangle| \leq \|\lambda_t\| \cdot \|\nabla G_t(\bar{\lambda})\| \leq M := ab(m+n)d$$

we can apply by the Azuma-Hoeffding inequality:

$$\mathsf{P}\left(-\frac{1}{T}\sum_{t=1}^{T}\langle\lambda_t, \nabla G_t(\bar{\lambda})\rangle \geq B\right) \leq \exp\left(-\frac{B^2T^2}{2\sum_{t=1}^{T} M^2}\right) = \exp\left(-\frac{1}{2}\frac{B^2}{M^2}T\right).$$

Using this inequality, we get

$$\mathsf{P}\left(\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t))) \geq \frac{A}{T^{1/4}}\right) \leq \exp\left(-\frac{A^2}{8M^2}\sqrt{T}\right) + \exp\left(-\frac{1}{512}\frac{A^4}{M^2\kappa^2 b^2 d^2}\right)$$

for sufficiently large A. This means that

$$\frac{1}{T}\sum_{t=1}^{T}(F_t(z_t^*) - F_t(\widetilde{z}_t(\lambda_t)) = O_{\mathsf{P}}\left(\frac{1}{T^{1/4}}\right).$$

To prove

$$\frac{1}{T}\sum_{t=1}^{T}\Delta\widetilde{z}_t(\lambda_t) = O_{\mathsf{P}}\left(\frac{1}{T^{1/4}}\right)$$

consider the representation

$$\frac{1}{T}\sum_{t=1}^{T}\Delta\widetilde{z}_t(\lambda_t) = \frac{1}{T}\sum_{t=1}^{T}\nabla G_t(\lambda_t) = \frac{1}{T}\sum_{t=1}^{T}\nabla G_t(\overline{\lambda})+\frac{1}{T}\sum_{t=1}^{T}(\nabla G_t(\lambda_t)-\nabla G_t(\overline{\lambda})).$$

Since $G_t(\overline{\lambda})$ are i.i.d. and $\mathsf{E}\nabla G_t(\overline{\lambda}) = 0$, from the central limit theorem it follows that

$$\frac{1}{T}\sum_{t=1}^{T}\nabla G_t(\overline{\lambda}) = O_{\mathsf{P}}\left(\frac{1}{\sqrt{T}}\right).$$

The estimate

$$\frac{1}{T}\sum_{t=1}^{T}(\nabla G_t(\lambda_t) - \nabla G_t(\overline{\lambda})) = O_{\mathsf{P}}\left(\frac{1}{T^{1/4}}\right)$$

is proved using the same technique as above. □

Computer experiments

$d = 1,\ m = 15,\ n = 15,\ a = 10,$

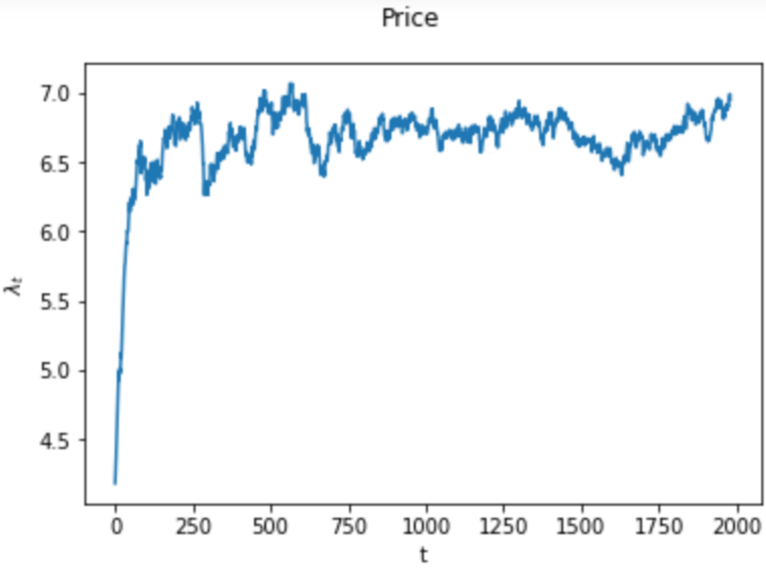
$$f_i(x_i) = \frac{A}{\alpha}(x_i + \varepsilon_1)^\alpha - \frac{A}{\alpha}\varepsilon_1^\alpha, \quad g_i(y_i) = \frac{B}{\beta}(y_i + \varepsilon_2)^\beta - \frac{B}{\beta}\varepsilon_2^\beta,$$

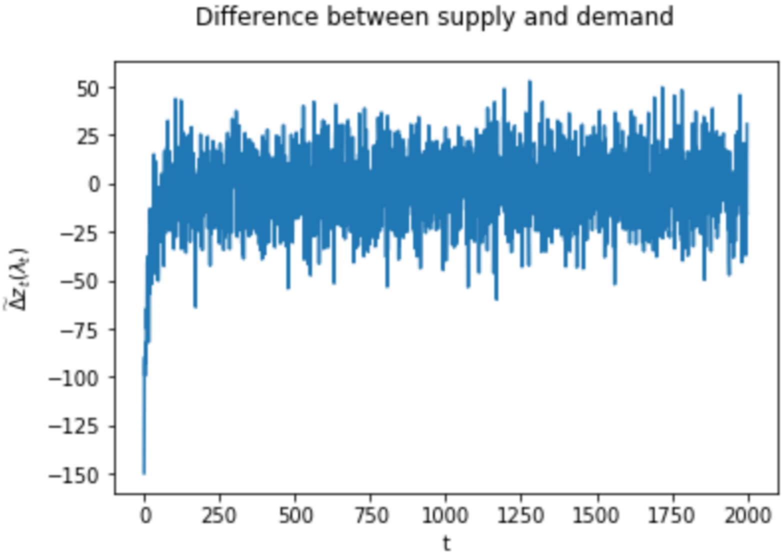
$$A \sim U(0, 15), \quad B \sim U(0, 10),$$

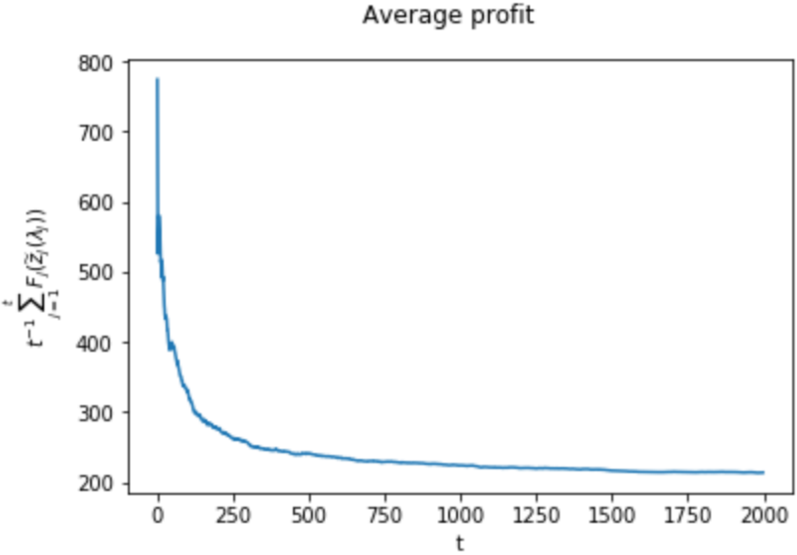
$$\varepsilon_1 \sim 0.1 + U(0, 1), \quad \varepsilon_2 \sim 0.1 + U(0, 1),$$

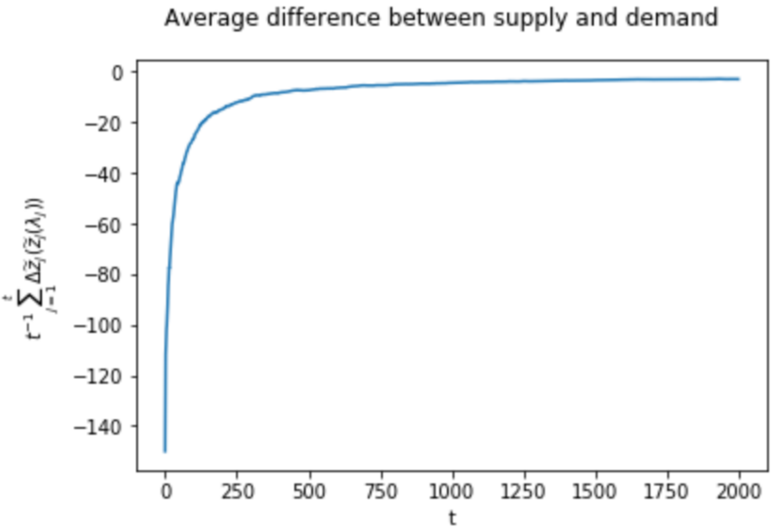
$$\alpha \sim 1 - U(0, 1), \quad \beta = 1 + U(0, 3),$$

$T = 2000.$


Price
7.0
6.5
6.0
5.5
5.0
4.5
λ_t
0
250
500
750
1000
1250
1500
1750
2000
t


Difference between supply and demand
50
25
0
−25
−50
−75
−100
−125
−150
0
250
500
750
1000
1250
1500
1750
2000
t
$\widetilde{\Delta z}_t(\lambda_t)$


Average profit
800
700
600
500
400
300
200
0
250
500
750
1000
1250
1500
1750
2000
t
$t^{-1}\sum_{j=1}^{t} F_j(\tilde{z}_j(\lambda_j))$


Average difference between supply and demand
0
−20
−40
−60
−80
−100
−120
−140
0
250
500
750
1000
1250
1500
1750
2000
t
$t^{-1}\sum_{j=1}^{t} \Delta\tilde{z}_j(\tilde{z}_j(\lambda_j))$

Agarwal, A. et al. (2006). "Algorithms for Portfolio Management Based on the Newton Method". In: *Proceedings of the 23rd International Conference on Machine Learning*. ICML '06, 9–16.

Azoury, K.S. and M.K. Warmuth (2001). "Relative loss bounds for on-line density estimation with the exponential family of distributions". In: *Machine Learning* 43.3, pp. 211–246.

Cover, T.M. (1991). "Universal Portfolios". In: *Mathematical Finance* 1.1, pp. 1–29. DOI: 10.1111/j.1467-9965.1991.tb00002.x.

Hazan, E. (2019). *Introduction to Online Convex Optimization*. arXiv: 1909.05207 [cs.LG].

Hazan, E. et al. (2006). "Logarithmic Regret Algorithms for Online Convex Optimization". In: *Learning Theory*. Ed. by G. Lugosi and H.U. Simon. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 499–513.

Helmbold, D.P. et al. (1998). "On-line portfolio selection using multiplicative updates". In: *Mathematical Finance* 8.4, pp. 325–347. DOI: 10.1111/1467-9965.00058.

Joulani, P., A. György, and C. Szepesvári (2020). "A modular analysis of adaptive (non-)convex optimization: Optimism, composite objectives, variance reduction, and variational bounds". In: *Theoretical Computer Science* 808, pp. 108–138.

- Kivinen, J. and M.K. Warmuth (1997). "Exponentiated gradient versus gradient descent for linear predictors". In: *Information and computation* 132.1, pp. 1–63. DOI: 10.1006/inco.1996.2612.
- Li, B. Li et al. (2015). "Moving average reversion strategy for on-line portfolio selection". In: *Artificial Intelligence* 222, pp. 104–123.
- Nesterov, Yu (1983). "A method of solving a convex programming problem with convergence rate $O(1/k^2)$". In: *Soviet Math. Dokl* 27.2, pp. 372–376.
- Orabona, F. (2020). *A Modern Introduction to Online Learning*. arXiv: 1912.13213 [cs.LG].
- Orabona, F. and D. Pál (2018). "Scale-free online learning". In: *Theoretical Computer Science* 716, pp. 50–69. DOI: 10.1016/j.tcs.2017.11.021.
- Orabona, Francesco and Dávid Pál (2016). "Coin Betting and Parameter-Free Online Learning". In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. NIPS'16. Barcelona, Spain, 577–585.
- Su, W., S. Boyd, and E.J. Candès (2016). "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights". In: *Journal of Machine Learning Research* 17.153, pp. 1–43.
- Vovk, V. (2001). "Competitive On-line Statistics". In: *International Statistical Review* 69.2, pp. 213–248.

Zinkevich, M. (2003). "Online convex programming and generalized infinitesimal gradient ascent". In: *Proceedings of the 20th international conference on machine learning (ICML-03)*, pp. 928–936.