Fast and Optimal Online Portfolio Selection

Dmitrii M. Ostrovskii

Joint work with

Rémi Jézéquel **Pierre Gaillard**

SICO 2022
June 16, 2022

Hero of the day

- The person from whom I learned the term "cross-term."
- Circa 2014 (upon my arrival to Grenoble to start a PhD):

A: *Dima, do you have a Facebook account?*
D: Erm, yes, but actually I d...
A: *Delete it!*

Dear Tolya, biz hundert un tsvantsig!!

(and then have a nice day!)

Outline

1. Online portfolio selection framework.
2. Cover's algorithm.
3. Our algorithm and main result.
4. Why it works.
5. Prospects.

Online portfolio selection

$\sim$30 years ago, Cover (1991) proposed a protocol for periodic investment:

Online portfolio selection

For each $t \in [T] := \{1, 2, ..., T\}$:

- Select $w_t \in \Delta_d$ // distribution of wealth over d assets chosen
- Receive $x_t \in \mathbb{R}_+^d$ // asset returns in this trading period revealed
- Suffer the loss $\ell_t(w_t) := -\log(x_t^\top w_t)$.

- **Motivation:** $x_t^\top w_t = e^{-\ell_t(w_t)}$ is the return of portfolio w_t chosen at t,

$$\prod_{t \leqslant T} x_t^\top w_t = \exp\left(-\sum_{t \leqslant T} \ell_t(w_t)\right)$$

is the cumulative return after T rounds, and its negative logarithm

$$\sum_{t \leqslant T} \ell_t(w_t).$$

is the **cumulative loss**.

Performance measure

- **Goal:** select w_t with large cumulative return, i.e. small cumulative loss.
- **Baseline:** the best **offline, static** selection
  - Also called the best "constantly-rebalanced portfolio" (CRP).

Regret of portfolio sequence $w_{1:T}$ given the market $x_{1:T}$:

$$\mathrm{Regret}_T(w_{1:T}|x_{1:T}) := \sum_{t \leqslant T} \ell_t(w_t) - \min_{w \in \Delta_d} \sum_{t \leqslant T} \ell_t(w).$$

Adversarial setup: must bound $\mathrm{Regret}_T(w_{1:T}|x_{1:T})$ uniformly over $x_{1:T}$:

$$\mathrm{Regret}_T(w_{1:T}|x_{1:T}) \leqslant f(T, d).$$

- Cover (1991) called any algorithm admitting such a bound **universal.**
- Naturally, we'd prefer a **sublinear regret:** $f(T, d) = o(T)$ for large T.

Cover's algorithm

Follow-the-leader:

$$w_{t+1} \in \operatorname*{Argmin}_{w \in \Delta_d} \sum_{\tau \leqslant t} \ell_\tau(w),$$

i.e. select the currently best CRP—for the market observed so far.

- **Fails:** for $(x_1, x_2, x_3, x_4, \ldots) = (e_1, e_2, e_1, e_2, \ldots)$ in $\mathbb{R}^2$ the regret is $\Omega(T)$.
- **Key trick:** robustify the procedure by exponential weighting:

Cover's algorithm (a.k.a. Exponentially Weighted Average Forecaster)

Consider the **distribution** on Δ_d with density

$$\phi_t(w) \propto \exp\left(-\sum_{\tau \leqslant t} \ell_\tau(w)\right), \quad w \in \Delta_d,$$

and play the averaged portfolio:

$$w_{t+1} = \mathbb{E}_{w \sim \phi_t}[w] \propto \int_{\Delta_d} w \exp\left(-\sum_{\tau \leqslant t} \ell_\tau(w)\right) dw.$$

Cover's algorithm (cont.)

Theorem (Cover, 1991)

For any realization $x_{1:T}$ of the market, Cover's algorithm has regret at most

$$d \log(eT).$$

Strong regret guarantee: not only sublinear in T, but near-constant.

- Doesn't depend on the magnitudes of x_t's and is linear in d.
- Unimprovable for $T \gtrsim d$ (Cesa-Bianchi and Lugosi, 2006).

Computationally prohibitive: requires integration in R^d (over density P_t).

- Sampling $\propto P_t$ can be used to approximate the integral/expectation.
- Kalai and Vempala (2002): implementation with $O(d^4 T^{14})$ runtime.
- Can be improved via better sampling, but this seems to be a dead end.

Open problem

Propose a **regret-optimal** and **computationally feasible** algorithm.

Progress timeline

Open problem

Propose a **regret-optimal** and **computationally feasible** algorithm.

Algorithm	Regret	Per-round runtime	Source
Cover's Algorithm	$d\log(T)$	$d^4 T^{14}$	Kalai and Vempala (2002)
Online Grad. Descent	$G\sqrt{Td}$	d	Zinkevich (2003)
Exponentiated Grad.	$G\sqrt{T\log(d)}$	d	Helmbold et al. (1998)
Online Newton Step	$Gd\log(T)$	d^3	Hazan et al. (2007)
Soft-Bayes	$\sqrt{Td\log(d)}$	d	Orseau et al. (2017)
Ada-BARRONS	$d^2\log^*(T)$	$d^{2.5}T$	Luo et al. (2018)
BISONS	$d^2\log^*(T)$	$\mathrm{poly}(d)$	Zimmert et al. (2022)
AdaMix+DONS	$d^2\log^*(T)$	$d^3\log^*(T)$	Mhammedi and Rakhlin (2022)
VB-FTRL	$d\log(T)$	$d^2 T$	**Our result**

Stepping stone: LB-FTRL

- Regularize the observed losses with the **log-barrier** of Δ_d, i.e. consider

$$\boxed{L_t(w) := \sum_{\tau \in [t]} \ell_t(w) \;+\; \lambda \sum_{i \in [d]} -\log(e_i^\top w)}$$

for some $\lambda > 0$. When $\lambda \geqslant 1$, this is a self-concordant (SC) barrier on Δ_d.

Log-Barrier FTRL

$$w_{t+1} \in \operatorname*{Argmin}_{w \in \Delta_d} L_t(w).$$

- Fast updates—only $O(d^2 T)$.
- Van Erven et al. (2020): $c\sqrt{dT\log(T)}$ regret when $\lambda \approx \sqrt{T/d} \gg 1$.
- They also **conjectured** that FTRL with $\lambda = O(1)$ is regret-optimal.
- Zimmert et al. (2022) **disproved** this by finding an adversary for which

$$\text{Regret}_T(w_{1:T}|x_{1:T}) \gtrsim \mathbf{2^d} \log(T) \log\log(T) \;\; \text{when} \;\; T \gtrsim \text{poly}(2^d).$$

Our algorithm: VB-FTRL

Perhaps we could "cure" LB-FTRL *without hurting its runtime?*

Key insight: it suffices to add to $L_t(w)$ a **volumetric** regularizer:

$$V_t(w) := \tfrac{1}{2} \log \det[\nabla^2 L_t(w)].$$

- $V_t(w)$ is the **volumetric barrier** for the domain in the FTL update:

$$\{w \in \Delta_d : \quad x_\tau^\top w > 0, \ \forall\, \tau \leqslant t\}.$$

- First proposed by Vaidya (1989) as a **self-concordant barrier** for LPs, with improved properties compared to the log-barrier—which is $L_t(w)$.

VB-FTRL – Follow-The-Regularized-Leader with the Volumetric Barrier:

$$w_{t+1} = \operatorname*{argmin}_{w \in \Delta_d} L_t(w) + \mu V_t(w) \quad \text{with some} \quad \lambda, \mu > 0.$$

- Results of Vaidya (1989) imply that $L_t + \mu V_t$ is **self-concordant** —informally, its 3rd derivatives are bounded in terms of the Hessian.
- Self-concordant functions are minimized well with **Newton's method.**

Main result

$$w_{t+1} = \operatorname*{argmin}_{w \in \Delta_d} L_t(w) + \mu V_t(w).$$

Theorem

For VB-FTRL with $\lambda = 16$ and $\mu = 6.5$ one has, uniformly over $x_{1:T}$, that

$$\mathrm{Regret}_T(\text{VB-FTRL}) \leqslant 26\, d \log(T + 16d).$$

- In other words, VB-FTRL is **regret-optimal** up to a constant factor.
- At the same time, VB-FTRL is **computationally cheap:**

Proposition

Newton's method run on $L_t(w) + \mu V_t(w)$ from w_t, converges in $\widetilde{O}(1)$ steps.

Computing a Newton step for $L_t(w) + \mu V_t(w)$ takes $O(d^2 T)$ when $T \gtrsim d$.

VB-FTRL as an approximation of Cover's algorithm

Besides its close form expression, P_{t-1} can also be seen as a Regularized Leader on probability distributions with entropic regularization,

$$P_{t-1} = \operatorname*{argmin}_{P \in \mathcal{P}(\Delta_d)} \left\{ \mathbb{E}_{w \sim P}[L_{t-1}(w)] - \mathcal{H}(P) \right\},$$

where $\mathcal{P}(\Delta_d)$ is the set of continuous probability distributions on the simplex and $\mathcal{H}(P) = -\mathbb{E}_{u\sim P} \log(P(u))$ is the differential entropy. We can remark that because L_{t-1} is self-concordant (see Appendix B.3 for details) and $\lambda \geq 1$, most of the mass of P_{t-1} will concentrated in the Dikin ellipsoid of L_{t-1} at w_t. In this ellipsoid, L_{t-1} is essentially quadratic i.e.

$$L_{t-1}(w) \simeq L_{t-1}(w_t) + \nabla L_{t-1}(w_t)^\top (w - w_t) + \frac{1}{2} \|w - w_t\|^2_{\nabla^2 L_{t-1}(w_t)}.$$

Therefore, P_{t-1} is close to a Gaussian distribution. This motivates optimizing over Gaussian distributions instead of all probability distributions making the optimization problem parametric.

$$\begin{aligned}
\theta_t, \Sigma_t &= \operatorname*{argmin}_{\theta \in \Delta_d, \Sigma \in \mathcal{S}_+^d} \left\{ \mathbb{E}_{w\sim\mathcal{N}(\theta,\Sigma)} L_{t-1}(w) - \mathcal{H}(\mathcal{N}(\theta,\Sigma)) \right\} \\
&= \operatorname*{argmin}_{\theta \in \Delta_d, \Sigma \in \mathcal{S}_+^d} \left\{ \mathbb{E}_{w\sim\mathcal{N}(\theta,\Sigma)} L_{t-1}(w) - \frac{1}{2} \log\det(\Sigma) \right\} \\
&\simeq \operatorname*{argmin}_{\theta \in \Delta_d, \Sigma \in \mathcal{S}_+^d} \left\{ L_{t-1}(\theta) + \frac{1}{2} \mathbb{E}_{w\sim\mathcal{N}(\theta,\Sigma)} \|w - \theta\|^2_{\nabla^2 L_{t-1}(\theta)} - \frac{1}{2} \log\det(\Sigma) \right\} \\
&= \operatorname*{argmin}_{\theta \in \Delta_d, \Sigma \in \mathcal{S}_+^d} \left\{ L_{t-1}(\theta) + \frac{1}{2} \mathrm{Tr}\left(\Sigma \nabla^2 L_{t-1}(\theta)\right) - \frac{1}{2} \log\det(\Sigma) \right\},
\end{aligned}$$

where we used that $\mathcal{H}(\mathcal{N}(\theta,\Sigma)) = \frac{1}{2}\log(|\Sigma|) + \mathrm{cst}$ and we have approximated L_{t-1} by its quadratic Taylor expansion in θ. In this approximated form, the optimization over Σ can be done in closed form and leads to

$$\theta_t \simeq \operatorname*{argmin}_{\theta \in \Delta_d} \left\{ L_{t-1}(\theta) + \frac{1}{2} \log\det(\nabla^2 L_{t-1}(\theta)) \right\}.$$

Thus, FTRL-VB (with $\mu = 1$) can be derived as an approximation of Universal Portfolio. To do so

Let $\widetilde{L}_t(w)$ be the objective minimized by VB-FTRL in each round:

$$\widetilde{L}_t(w) := L_t(w) + \mu V_t(w).$$

Proof. We shall specify the values of λ, μ in the final step of the proof; for now we consider the general case with $\lambda, \mu > 0$. We first observe that the regret $\mathsf{Regret} = \mathsf{Regret}(w_{1:T}, x_{1:T})$ admits the following decomposition:

$$\begin{aligned}
\mathsf{Regret} &= \sum_{t\in[T]} \Big[\ell_t(w_t) + \min_{w\in\Delta_d} \widetilde{L}_{t-1}(w) - \min_{w\in\Delta_d} \widetilde{L}_t(w)\Big] + \min_{w\in\Delta_d} \widetilde{L}_T(w) - \min_{w\in\Delta_d} \widetilde{L}_0(w) - \min_{w\in\Delta_d} \sum_{t\in[T]} \ell_t(w) \\
&\leqslant \sum_{t\in[T]} \Big[\ell_t(w_t) + \min_{w\in\Delta_d} \widetilde{L}_{t-1}(w) - \min_{w\in\Delta_d} \widetilde{L}_t(w)\Big] + \min_{w\in\Delta_d} \widetilde{L}_T(w) - \min_{w\in\Delta_d} \sum_{t\in[T]} \ell_t(w) \\
&= \sum_{t\in[T]} \underbrace{[\widetilde{L}_t(w_t) - \widetilde{L}_t(w_{t+1})]}_{\mathsf{Lagged}(t)} + \sum_{t\in[T]} \underbrace{\mu\,[V_{t-1}(w_t) - V_t(w_t)]}_{\mathsf{Offset}(t)} + \min_{w\in\Delta_d} \widetilde{L}_T(w) - \min_{w\in\Delta_d} \sum_{t\in[T]} \ell_t(w). \qquad (1)
\end{aligned}$$

Here the first identity is by telescoping, and the inequality is due to $\widetilde{L}_0(w) = \lambda R(w)$ being non-negative for any $w \in \Delta_d$. Our plan is as follows: first, we shall estimate the final term in the right-hand side of (1) as

$$\mathsf{Bias}\ \left[:= \min_{w\in\Delta_d} \widetilde{L}_T(w) - \min_{w\in\Delta_d} \sum_{t\in[T]} \ell_t(w)\right] \leqslant \left(\frac{3\mu}{2} + \lambda\right) d \log(T + \lambda d). \qquad (2)$$

Then we shall prove that $\mathsf{Offset}(t) + \mathsf{Lagged}(t) \leqslant 0$ is maintained for all $t \in [T]$, whence the theorem will follow.

Perspectives

(?) Acceleration from $O(d^2 T)$ to $O(d^3)$.

(?) "Quantum" portfolios (measurement of the quantum state).

(?) Other online learning problems with intractable near-optimal strategies, e.g. **online linear optimization with bandit feedback**.

Thanks!

References I

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2):169–192, 2007.

David P. Helmbold, Robert E. Schapire, Yoram Singer, and Manfred K. Warmuth. On-line portfolio selection using multiplicative updates. *Mathematical Finance*, 8 (4):325–347, 1998.

Adam T. Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, pages 423–440, 2002.

Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. In *Proceedings of the 33rd international Conference on Neural Information Processing Systems*, volume 31, pages 8245–8255, 2018.

Zakaria Mhammedi and Alexander Rakhlin. Damped online Newton step for portfolio selection. *arXiv preprint arXiv:2202.07574*, 2022.

References II

Laurent Orseau, Tor Lattimore, and Shane Legg. Soft-bayes: Prod for mixtures of experts with log-loss. In *International Conference on Algorithmic Learning Theory*, pages 372–399. PMLR, 2017.

Pravin M. Vaidya. A new algorithm for minimizing convex functions over convex sets. In *30th Annual Symposium on Foundations of Computer Science*, pages 338–343. IEEE Computer Society, 1989.

Tim Van Erven, Dirk Van der Hoeven, Wojciech Kotłowski, and Wouter M. Koolen. Open problem: Fast and optimal online portfolio selection. In *Proceedings of the 33rd Conference On Learning Theory*, pages 3864–3869. PMLR, 2020.

Julian Zimmert, Naman Agarwal, and Satyen Kale. Pushing the efficiency-regret Pareto frontier for online learning of portfolios and quantum states. *arXiv preprint arXiv:2202.02765*, 2022.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*, pages 928–936, 2003.