

Open Problem: Fast and Optimal Online Portfolio Selection

Tim van Erven

and **Dirk van der Hoeven**

Mathematical Institute, Leiden University, the Netherlands

TIM@TIMVANERVEN.NL

DIRK@DIRKVANDERHOEVEN.COM

Wojciech Kotłowski

Poznan University of Technology, Poland

KOTLOW@GMAIL.COM

Wouter M. Koolen

Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

WMKOOLEN@CWI.NL

Editors: Jacob Abernethy and Shivani Agarwal

Abstract

Online portfolio selection has received much attention in the COLT community since its introduction by Cover, but all state-of-the-art methods fall short in at least one of the following ways: they are either i) computationally infeasible; or ii) they do not guarantee optimal regret; or iii) they assume the gradients are bounded, which is unnecessary and cannot be guaranteed. We are interested in a natural follow-the-regularized-leader (FTRL) approach based on the log barrier regularizer, which is computationally feasible. The open problem we put before the community is to formally prove whether this approach achieves the optimal regret. Resolving this question will likely lead to new techniques to analyse FTRL algorithms. There are also interesting technical connections to self-concordance, which has previously been used in the context of bandit convex optimization.

1. Introduction

Online portfolio selection (Cover, 1991) may be viewed as an instance of online convex optimization (OCO) (Hazan et al., 2016): in each of $t = 1, \dots, T$ rounds, a learner has to make a prediction \mathbf{w}_t in a convex domain \mathcal{W} before observing a convex loss function $f_t : \mathcal{W} \rightarrow \mathbb{R}$. The goal is to obtain a guaranteed bound on the regret $\text{Regret}_T = \sum_{t=1}^T f_t(\mathbf{w}_t) - \min_{\mathbf{w} \in \mathcal{W}} \sum_{t=1}^T f_t(\mathbf{w})$ that holds for any possible sequence of loss functions f_t . Online portfolio selection corresponds to the special case that the domain $\mathcal{W} = \{\mathbf{w} \in \mathbb{R}_+^d \mid \sum_{i=1}^d w_i = 1\}$ is the probability simplex and the loss functions are restricted to be of the form $f_t(\mathbf{w}) = -\ln(\mathbf{w}^\top \mathbf{x}_t)$ for vectors $\mathbf{x}_t \in \mathbb{R}_+^d$. It was introduced by Cover (1991) with the interpretation that $x_{t,i}$ represents the factor by which the value of an asset $i \in \{1, \dots, d\}$ grows in round t and $w_{t,i}$ represents the fraction of our capital we re-invest in asset i in round t . The factor by which our initial capital grows over T rounds then becomes $\prod_{t=1}^T \mathbf{w}_t^\top \mathbf{x}_t = e^{-\sum_{t=1}^T f_t(\mathbf{w}_t)}$. An alternative interpretation in terms of mixture learning is given by Orseau et al. (2017).

For an extensive survey of online portfolio selection we refer to Li and Hoi (2014). Here we review only the results that are most relevant to our open problem. Cover (1991); Cover and Ordentlich (1996) show that the best possible guarantee on the regret is of order $\text{Regret}_T = O(d \ln T)$ and that this is achieved by choosing \mathbf{w}_{t+1} as the mean of a continuous exponential weights distribution $dP_{t+1}(\mathbf{w}) \propto e^{-\sum_{s=1}^t f_s(\mathbf{w})} d\pi(\mathbf{w})$ with Dirichlet-prior π (and learning rate $\eta = 1$). Unfortunately, this approach has a run-time of order $O(T^d)$, which scales exponentially in the number

of assets d , and is therefore computationally infeasible when d exceeds, say, 3. A sampling-based implementation by [Kalai and Vempala \(2002\)](#) greatly improves the run-time to $\tilde{O}(T^4(T+d)d^2)$, but even this is still infeasible already for modest d and T .

As shown in [Table 1](#), much faster algorithms are available, but they either do not achieve the optimal regret or they assume that the gradients are uniformly bounded by a *known* bound G : $\|\nabla f_t(\mathbf{w}_t)\|_2 \leq G$, and the bounds deteriorate rapidly when G is large. Bounding the gradients is very restrictive: we either need to (i) assume that the asset prices do not fluctuate too rapidly, which defeats the purpose of using adversarial online learning; or (ii) we need to allocate a minimum amount of capital $w_{t,i} \geq \alpha$ to each asset, which means we cannot drop any poorly performing assets from our portfolio.

Table 1: Overview of achievable trade-offs between regret and run-time

Method	Regret	Run-time	Assumes Bounded Gradients	References
Universal Portfolio	$O(d \ln(T))$	$\tilde{O}(T^4(T+d)d^2)$	No	(Cover and Ordentlich, 1996; Kalai and Vempala, 2002)
Online Newton Step	$O(Gd \ln(T))$	$O(d^3T)$	Yes	(Agarwal et al., 2006; Hazan et al., 2007; Hazan and Kale, 2015)
Exponentiated Gradient	$O(G\sqrt{T \ln(d)})$	$O(dT)$	Yes	Helmbold et al. (1998)
Gradient Descent	$O(G\sqrt{dT})$	$O(dT)$	Yes	Zinkevich (2003)
Soft-Bayes	$O(\sqrt{dT \ln(d)})$	$O(dT)$	No	Orseau et al. (2017)
Ada-BARRONS	$O(d^2 \ln^4(T))$	$O(d^{2.5}T^2)$	No	Luo et al. (2018)
FTRL	?	$O(d^2T^2)$	No	Agarwal and Hazan (2005)

We are interested in a natural follow-the-regularized-leader algorithm, previously proposed by [Agarwal and Hazan \(2005\)](#):

$$\mathbf{w}_{t+1} = \arg \min_{\mathbf{w} \in \mathcal{W}} \left\{ \sum_{s=1}^t f_s(\mathbf{w}) + \lambda \sum_{i=1}^d -\ln w_i \right\} \quad (1)$$

for some $\lambda > 0$. The regularizer $R(\mathbf{w}) = \sum_{i=1}^d -\ln w_i$ is a self-concordant barrier function ([Abernethy et al., 2008](#)) that is the log barrier for the positive orthant ([Foster et al., 2016](#)) and has a natural interpretation as adding d extra rounds in which \mathbf{x} equals $\mathbf{e}_1, \dots, \mathbf{e}_d$.

The optimization problem (1) can be solved to machine precision in $O(d^2t)$ steps using Newton’s method, so a naive implementation in which we solve the optimization problem independently for each round would already lead to a total run-time of $O(d^2T^2)$, which is computationally feasible for practical values of d and T . One might further hope that sharing calculations between rounds or solving (1) approximately may lead to additional speed-ups, similar to those obtained for FTRL with linear losses by [Abernethy et al. \(2008\)](#). Thus the method is computationally feasible, at least for an interesting range of d and T . The open problem we now pose is whether it is also worst-case optimal in terms of regret:

Open Problem: *Does the FTRL algorithm (1) guarantee the optimal regret $O(d \log T)$ without further assumptions like bounded gradients?*

Our motivation is twofold: efficient algorithms for portfolio selection (and beyond) are desirable, and FTRL is the simplest natural candidate. In addition, our current inability to analyse it highlights frustrating blind spots in our FTRL toolbox, which solving this problem will need to address.

Agarwal and Hazan (2005) already prove $O(G^2 d \ln(dT))$ regret when the gradients are bounded, but we believe that the bound should not depend on G at all. It seems that the key difficulty in analyzing the regret is to control the sum of so-called local norms of the gradients. As we will discuss below, this is possible at least in several encouraging special cases.

2. Technical Discussion

It is convenient to reparametrize by $\mathbf{v} \in \mathbb{R}_+^{d-1}$ such that $\sum_{i=1}^{d-1} v_i \leq 1$, obtaining $\mathbf{w}_t = A\mathbf{v}_t + \mathbf{b}$ for $A = \begin{pmatrix} I \\ -\mathbf{1}^\top \end{pmatrix}$, and $\mathbf{b} = \mathbf{e}_d$. With some abuse of notation, we will also write $f_t(\mathbf{v})$ for $f_t(A\mathbf{v} + \mathbf{b})$ and $R(\mathbf{v})$ for $R(A\mathbf{v} + \mathbf{b})$. Then the criterion being minimized is

$$\phi_T(\mathbf{v}) = \sum_{t=1}^T f_t(\mathbf{v}) + \lambda R(\mathbf{v}).$$

As the loss is 1-exp-concave, we have $\nabla^2 f_t(\mathbf{v}) \succeq \nabla f_t(\mathbf{v}) \nabla f_t(\mathbf{v})^\top$ (Bubeck, 2015, pp. 324–325). In fact, this holds with equality in the present case:

$$\nabla f_t(\mathbf{v}) = \frac{-A^\top \mathbf{x}_t}{(A\mathbf{v} + \mathbf{b})^\top \mathbf{x}_t}, \quad \nabla^2 f_t(\mathbf{v}) = \frac{A^\top \mathbf{x}_t \mathbf{x}_t^\top A}{((A\mathbf{v} + \mathbf{b})^\top \mathbf{x}_t)^2} = \nabla f_t(\mathbf{v}) \nabla f_t(\mathbf{v})^\top.$$

2.1. Regret Bounded by Local Norms via Self-concordance

We observe that both the losses f_t and the regularizer R are self-concordant functions (Abernethy et al., 2008). Assume for simplicity that $\lambda \geq 1$, in which case ϕ_T is a sum of self-concordant functions and hence also self-concordant. Like Abernethy et al. (2008), define the local norms $\|\mathbf{g}\|_t = \sqrt{\mathbf{g}^\top \nabla^{-2} \phi_t(\mathbf{v}_t) \mathbf{g}}$. The gradients are always bounded in these local norms:

$$\|\nabla f_t(\mathbf{v}_t)\|_t^2 \leq \nabla f_t(\mathbf{v}_t)^\top (\nabla f_t(\mathbf{v}_t) \nabla f_t(\mathbf{v}_t)^\top + \lambda \nabla^2 R(\mathbf{v}_t))^{-1} \nabla f_t(\mathbf{v}_t) \leq \frac{1}{1 + \lambda}, \quad (2)$$

where the last inequality follows because $\nabla^2 R(\mathbf{v}) \succeq \nabla f_t(\mathbf{v}) \nabla f_t(\mathbf{v})^\top$ for all \mathbf{v} (proof omitted).

Lemma 1 *For $\lambda \geq 1$, the regret is bounded in terms of the local norms:*

$$\text{Regret}_T = O\left(\lambda d \log T + \sum_{t=1}^T \|\nabla f_t(\mathbf{v}_t)\|_t^2\right).$$

Proof sketch: Let $\mathbf{v}^* \in \arg \min_{\mathbf{v}} \sum_{t=1}^T f_t(\mathbf{v})$. Then

$$\text{Regret}_T = \phi_T(\mathbf{v}_{T+1}) - \lambda R(\mathbf{v}_1) - \sum_{t=1}^T f_t(\mathbf{v}^*) + \sum_{t=1}^T (\phi_t(\mathbf{v}_t) - \phi_t(\mathbf{v}_{t+1})).$$

The result follows by bounding $\phi_T(\mathbf{v}_{T+1}) \leq \phi_T((1 - \frac{1}{2T})\mathbf{v}^* + \frac{1}{2T}\mathbf{v}_1) \leq \lambda R(\mathbf{v}_1) + \sum_{t=1}^T f_t(\mathbf{v}^*) + \lambda d \log(2T) + 1$ and using (2.16) of Nemirovski (2004) for the self-concordant function ϕ_t to bound $\phi_t(\mathbf{v}_t) - \phi_t(\mathbf{v}_{t+1}) \leq C \|\nabla \phi_t(\mathbf{v}_t)\|_t^2 = C \|\nabla f_t(\mathbf{v}_t)\|_t^2$ for a suitable constant $C > 0$ that can be made independent of λ for $\lambda \geq 1$. Here the fact that $\nabla \phi_t(\mathbf{v}_t) = \nabla f_t(\mathbf{v}_t)$ followed because $\phi_t(\mathbf{v}_t) = f_t(\mathbf{v}_t) + \phi_{t-1}(\mathbf{v}_t)$ and \mathbf{v}_t minimizes ϕ_{t-1} in the interior of \mathcal{W} . \blacksquare

Combining Lemma 1 with (2), we immediately see that the regret is bounded by

$$\text{Regret}_T = O(\sqrt{dT \log T}) \quad \text{for } \lambda \approx \sqrt{\frac{T}{d \log T}},$$

but if we hope to get the optimal rate, we need to use constant λ , so this is what we will assume from now on. Below we list several promising corollaries of Lemma 1.

2.2. Assuming Bounded Gradients

Suppose that, for some reason, the gradients with respect to \mathbf{w} (not \mathbf{v} !) are bounded: $\|\nabla f_t(\mathbf{w}_t)\|_2 = \|\frac{-\mathbf{x}_t}{\mathbf{w}_t^\top \mathbf{x}_t}\|_2 \leq G$. Then, abbreviating $\mathbf{y}_t = A^\top \mathbf{x}_t / \|\mathbf{x}_t\|_2$, we can use that $\nabla^2 f_t(\mathbf{v}_t) \succeq \frac{A^\top \mathbf{x}_t \mathbf{x}_t^\top A}{\|\mathbf{x}_t\|_\infty^2} \succeq \frac{A^\top \mathbf{x}_t \mathbf{x}_t^\top A}{\|\mathbf{x}_t\|_2^2} = \mathbf{y}_t \mathbf{y}_t^\top$ to get

$$\sum_{t=1}^T \|\nabla f_t(\mathbf{v}_t)\|_t^2 \leq G^2 \sum_{t=1}^T \|\mathbf{y}_t\|_t^2 \leq G^2 \sum_{t=1}^T \mathbf{y}_t^\top \left(\sum_{s=1}^t \mathbf{y}_s \mathbf{y}_s^\top + \lambda A^\top A \right)^{-1} \mathbf{y}_t = O(G^2 d \log T),$$

where the last step follows analogously to Hazan et al. (2007, Lemma 11) and using that $\det(A^\top A) = d$ by Sylvester's determinant theorem. This gives the optimal rate if G is small.

2.3. Source Coding and \mathbf{x}_t in Finite Set

We call the case that $\mathbf{x}_t \in \{e_1, \dots, e_d\}$ the source coding setting. This case is easy to analyse, because \mathbf{w}_t has a simple closed-form solution that coincides with Cover's universal portfolio algorithm. More generally, let us assume that \mathbf{x}_t takes values in some finite set \mathcal{X} of size k , so $k = d$ in the source coding setting, and let $n_t(\mathbf{x})$ denote the number of times that $\mathbf{x}_s = \mathbf{x}$ for $s \leq t$. Then

$$\begin{aligned} \sum_{t=1}^T \|\nabla f_t(\mathbf{v}_t)\|_t^2 &\leq \sum_{t=1}^T \nabla f_t(\mathbf{v}_t)^\top (n_t(\mathbf{x}_t) \nabla f_t(\mathbf{v}_t) \nabla f_t(\mathbf{v}_t)^\top + \lambda \nabla^2 R(\mathbf{v}_t))^{-1} \nabla f_t(\mathbf{v}_t) \\ &\leq \sum_{t=1}^T \frac{1}{n_t(\mathbf{x}_t) + \lambda} = \sum_{\mathbf{x} \in \mathcal{X}} \sum_{j=1}^{n_T(\mathbf{x})} \frac{1}{j + \lambda} = O\left(\sum_{\mathbf{x} \in \mathcal{X}} \log n_T(\mathbf{x})\right) = O(k \log T). \end{aligned}$$

In particular, algorithm (1) achieves the optimal rate in the source coding setting.

2.4. A (Suboptimal) General Bound without Bounded Gradients

Since $R(\mathbf{v})$ is a barrier, it should be the case that $w_{t,i} \geq C/t$ for some constant $C > 0$. We may therefore cover the effective domain of \mathbf{w}_t by $m = O((\log T)^d)$ sets B_1, \dots, B_m such that $\mathbf{w}^\top \mathbf{x} \leq 2\mathbf{u}^\top \mathbf{x}$ for all $\mathbf{w}, \mathbf{u} \in B_i$. It follows that

$$\begin{aligned} \sum_{t=1}^T \|\nabla f_t(\mathbf{v}_t)\|_t^2 &\leq \sum_{i=1}^m \sum_{t: \mathbf{v}_t \in B_i} \nabla f_t(\mathbf{v}_t)^\top \left(\sum_{s \leq t: \mathbf{v}_s \in B_i} \nabla f_s(\mathbf{v}_s) \nabla f_s(\mathbf{v}_s)^\top + \lambda \nabla^2 R(\mathbf{v}_t) \right)^{-1} \nabla f_t(\mathbf{v}_t) \\ &\leq 4 \sum_{i=1}^m \sum_{t: \mathbf{v}_t \in B_i} \nabla f_t(\mathbf{v}_t)^\top \left(\sum_{s \leq t: \mathbf{v}_s \in B_i} \nabla f_s(\mathbf{v}_s) \nabla f_s(\mathbf{v}_s)^\top + \lambda A^\top A \right)^{-1} \nabla f_t(\mathbf{v}_t) \\ &= O(md \log T) = O(d(\log T)^{d+1}), \end{aligned}$$

where the first equality follows like Lemma 11 of Hazan et al. (2007) with $\|\nabla f_t(\mathbf{v}_t)\| \leq t/C$. This of course has wildly suboptimal dependence in d , but shows near-optimal regret for very small d .

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO): grant numbers VI.Vidi.192.095 for Van Erven and TOP2EW.15.211 for Van der Hoeven. Kotłowski was supported by the Polish National Science Centre (grant no. 2016/22/E/ST6/00299).

References

- Jacob Abernethy, Elad Hazan, and Alexander Rakhlin. Competing in the dark: An efficient algorithm for bandit linear optimization. In *COLT*, pages 263–274, 2008.
- Amit Agarwal and Elad Hazan. Efficient algorithms for online game playing and universal portfolio management. *ECCC, TR06-033*, 2005.
- Amit Agarwal, Elad Hazan, Satyen Kale, and Robert E Schapire. Algorithms for portfolio management based on the newton method. In *Proceedings of the 23rd international conference on Machine learning*, pages 9–16, 2006.
- Sébastien Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends in Machine Learning*, 8(3–4):231–357, 2015.
- T. M. Cover and E. Ordentlich. Universal portfolios with side information. *IEEE Transactions on Information Theory*, 42(2):348–363, 1996.
- Thomas M Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- Dylan J Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Eva Tardos. Learning in games: Robustness of fast convergence. In *Advances in Neural Information Processing Systems 29*, pages 4734–4742. 2016.
- Elad Hazan and Satyen Kale. An online portfolio selection algorithm with regret logarithmic in price variation. *Mathematical Finance*, 25(2):288–310, 2015.
- Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- Elad Hazan et al. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.
- 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 Kalai and Santosh Vempala. Efficient algorithms for universal portfolios. *Journal of Machine Learning Research*, 3(Nov):423–440, 2002.
- Bin Li and Steven CH Hoi. Online portfolio selection: A survey. *ACM Computing Surveys (CSUR)*, 46(3):1–36, 2014.
- Haipeng Luo, Chen-Yu Wei, and Kai Zheng. Efficient online portfolio with logarithmic regret. In *Advances in Neural Information Processing Systems*, pages 8235–8245, 2018.

Arkadi Nemirovski. Lecture notes: Interior point polynomial time methods in convex programming. *Spring Semester*, 2004.

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, 2017.

Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. *Proceedings of the Twenty International Conference on Machine Learning (ICML'03)*, (T. Fawcett and N. Mishra, eds.), pp. 928-936, AAAI Press, 2003.