The many faces of online learning

Hoeven, D. van der

Author: Hoeven, D. van der
Title: The many faces of online learning
Issue Date: 2021-03-4

CHAPTER 7

Open Problem: Fast and Optimal Online Portfolio Selection

This chapter is based on Van Erven, T., Van der Hoeven, D., Kotłowski, W., and Koolen, W. M. (2020b). Open problem: Fast and optimal online portfolio selection. In *Proceedings of the 33rd Annual Conference on Learning Theory (COLT)*, pages 3864–3869.[1]

Abstract

Online portfolio selection has received much attention 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]The author of this dissertation performed the following tasks: co-deriving the theoretical results and co-writing the paper.

7.1 Introduction

Online portfolio selection (Cover, 1991) may be viewed as an instance of online convex optimization (OCO) (Hazan et al., 2016): in each of $t = 1, \ldots, T$ rounds, a learner has to make a prediction $\boldsymbol{w}_t$ in a convex domain $\mathcal{W}$ before observing a convex loss function $f_t : \mathcal{W} \to \mathbb{R}$. The goal is to obtain a guaranteed bound on the regret $\mathcal{R}_T = \sum_{t=1}^T f_t(\boldsymbol{w}_t) - \min_{\boldsymbol{w}\in\mathcal{W}} \sum_{t=1}^T f_t(\boldsymbol{w})$ that holds for any possible sequence of loss functions f_t. Online portfolio selection corresponds to the special case that the domain $\mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}_+^d \mid \sum_{i=1}^d w_i = 1\}$ is the probability simplex and the loss functions are restricted to be of the form $f_t(\boldsymbol{w}) = -\ln(\boldsymbol{w}^\intercal \boldsymbol{x}_t)$ for vectors $\boldsymbol{x}_t \in \mathbb{R}_+^d$. It was introduced by Cover (1991) with the interpretation that $x_{t,i}$ represents the factor by which the value of an asset $i \in \{1, \ldots, d\}$ 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 $\prod_{t=1}^T \boldsymbol{w}_t^\intercal \boldsymbol{x}_t = e^{-\sum_{t=1} f_t(\boldsymbol{w}_t)}$. 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 $\mathcal{R}_T = O(d \ln T)$ and that this is achieved by choosing $\boldsymbol{w}_{t+1}$ as the mean of a continuous exponential weights distribution $\mathrm{d}P_{t+1}(\boldsymbol{w}) \propto e^{-\sum_{s=1}^t f_s(\boldsymbol{w})} \mathrm{d}\pi(\boldsymbol{w})$ 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 7.1, 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: $\|\nabla f_t(\boldsymbol{w}_t)\|_2 \le G$, 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 $w_{t,i} \ge \alpha$ to each asset, which means we cannot drop any poorly performing assets from our portfolio.

We are interested in a natural follow-the-regularized-leader algorithm, previously

Table 7.1: 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)

proposed by Agarwal and Hazan (2005):

$$\boldsymbol{w}_{t+1} = \operatorname*{arg\,min}_{\boldsymbol{w}\in\mathcal{W}} \left\{ \sum_{s=1}^{t} f_s(\boldsymbol{w}) + \lambda \sum_{i=1}^{d} -\ln w_i \right\} \tag{7.1.1}$$

for some $\lambda > 0$. The regularizer $R(\boldsymbol{w}) = \sum_{i=1}^{d} -\ln w_i$ is a self-concordant barrier function (Nesterov and Nemirovskii, 1994) that is the log barrier for the positive orthant and has a natural interpretation as adding d extra rounds in which $\boldsymbol{x}$ equals $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$.

The optimization problem (7.1.1) 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 (7.1.1) 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* (7.1.1) *guarantee the optimal regret* $O(d\ln T)$ *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.

7.2 Technical Discussion

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

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

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

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

7.2.1 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 $\|\boldsymbol{g}\|_t = \sqrt{\boldsymbol{g}^\intercal \nabla^{-2} \phi_t(\boldsymbol{v}_t) \boldsymbol{g}}$. By Lemma 28 below we know that the gradients are always bounded in these local norms.

Lemma 28. $\|\nabla f_t(\boldsymbol{v}_t)\|_t^2 \leq \frac{1}{\lambda+1}$

Proof. We start by observing that

$$\begin{aligned}
\|\nabla f_t(\boldsymbol{v}_t)\|_t^2 \leq & \nabla f_t(\boldsymbol{v}_t)^\intercal (\nabla f_t(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t)^\intercal + \lambda \nabla^2 R(\boldsymbol{v}_t))^{-1} \nabla f_t(\boldsymbol{v}_t) \\
= & \lambda^{-1} \|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2 - \frac{\lambda^{-2} \|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^4}{1 + \lambda^{-1} \|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2} \\
= & \frac{\|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2}{\lambda + \|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2},
\end{aligned}$$

where $\|\boldsymbol{g}\|_{R(\boldsymbol{v}_t)} = \sqrt{\boldsymbol{g}^\intercal \nabla^{-2} R(\boldsymbol{v}_t)\boldsymbol{g}}$ and the first equality follows from the Sherman-Morrison formula. Note that $\nabla^2 R(\boldsymbol{v}_t)$ is positive definite, so $\nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^2 R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t) = 0$ only when $\nabla f_t(\boldsymbol{v}_t) = \mathbf{0}$, for which the result holds. If $\nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^2 R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t) > 0$, because $\nabla^2 R(\boldsymbol{v}) \succeq \nabla f_t(\boldsymbol{v}) \nabla f_t(\boldsymbol{v})^\intercal$ for all $\boldsymbol{v}$ by Lemma 29 below we have

$$\begin{aligned}
&\nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^{-2} R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^2 R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t) \\
&\leq \nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^2 R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t)
\end{aligned}$$

and thus $\nabla f_t(\boldsymbol{v}_t)^\intercal \nabla^{-2} R(\boldsymbol{v}_t) \nabla f_t(\boldsymbol{v}_t) \leq 1$. Using that $s(x) = x/(\lambda + x)$ is increasing for $x > -\lambda$ we conclude that the gradients are indeed bounded in the local norms:

$$\|\nabla f_t(\boldsymbol{v}_t)\|_t^2 \leq \frac{\|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2}{\lambda + \|\nabla f_t(\boldsymbol{v}_t)\|_{R(\boldsymbol{v}_t)}^2} \leq \frac{1}{\lambda + 1}. \tag{7.2.1}$$

□

Lemma 29. $\nabla^2 R(\boldsymbol{v}) \succeq \nabla f_t(\boldsymbol{v}) \nabla f_t(\boldsymbol{v})^\intercal$ *for all* $\boldsymbol{v}$*.*

Proof. We need to show that, for all $\boldsymbol{x}$ and $\boldsymbol{w}$:

$$A^\intercal \Big(\sum_{i=1}^d \frac{\boldsymbol{e}_i \boldsymbol{e}_i^\intercal}{\big(\boldsymbol{w}^\intercal \boldsymbol{e}_i\big)^2} \Big) A \succeq \frac{A^\intercal \boldsymbol{x}\boldsymbol{x}^\intercal A}{\big(\boldsymbol{w}^\intercal \boldsymbol{x}\big)^2}.$$

It is sufficient to show that:

$$\sum_{i=1}^d \frac{\boldsymbol{e}_i \boldsymbol{e}_i^\intercal}{\big(\boldsymbol{w}^\intercal \boldsymbol{e}_i\big)^2} \succeq \frac{\boldsymbol{x}\boldsymbol{x}^\intercal}{\big(\boldsymbol{w}^\intercal \boldsymbol{x}\big)^2}.$$

Both sides are positive semi-definite. The right-hand side is rank 1, with eigenvector $\boldsymbol{x}$. Hence it is sufficient to show that

$$\begin{aligned}
\boldsymbol{x}^\intercal \Big(\sum_{i=1}^d \frac{\boldsymbol{e}_i \boldsymbol{e}_i^\intercal}{\big(\boldsymbol{w}^\intercal \boldsymbol{e}_i\big)^2} \Big) \boldsymbol{x} &\geq \boldsymbol{x}^\intercal \Big(\frac{\boldsymbol{x}\boldsymbol{x}^\intercal}{\big(\boldsymbol{w}^\intercal \boldsymbol{x}\big)^2} \Big) \boldsymbol{x} \\
\sum_{i=1}^d \frac{x_i^2}{w_i^2} &\geq \frac{\|\boldsymbol{x}\|_2^4}{\big(\boldsymbol{w}^\intercal \boldsymbol{x}\big)^2} \\
\big(\boldsymbol{w}^\intercal \boldsymbol{x}\big)^2 \sum_{i=1}^d \frac{x_i^2}{w_i^2} &\geq \|\boldsymbol{x}\|_2^4 \\
\|\boldsymbol{y}\|_1^2 \|\boldsymbol{z}\|_2^2 &\geq (\boldsymbol{y}^\intercal \boldsymbol{z})^2 \\
\|\boldsymbol{y}\|_1 \|\boldsymbol{z}\|_2 &\geq \boldsymbol{y}^\intercal \boldsymbol{z}
\end{aligned}$$

where $\boldsymbol{y} = (y_1, \ldots, y_d)$ for $y_i = w_i x_i$, $\boldsymbol{z} = (z_1, \ldots, z_d)$ for $z_i = x_i / w_i$, and we are using that $w_i, x_i \geq 0$. The result then follows upon observing that $\|\boldsymbol{z}\|_1 \geq \|\boldsymbol{z}\|_2$, and applying the Cauchy-Schwarz inequality. $\square$

To bound the regret the following Lemma is useful.

Lemma 30. *For $\lambda \geq \frac{5}{4}$, the regret is bounded in terms of the local norms:*

$$\mathcal{R}_T \leq \lambda d \ln(2T) + 1 + \sum_{t=1}^{T} \|\nabla f_t(\boldsymbol{v}_t)\|_t^2.$$

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

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

We start by bounding

$$\begin{aligned}
\phi_T(\boldsymbol{v}_{T+1}) - \lambda R(\boldsymbol{v}_1) \leq & \phi_T\big((1 - \tfrac{1}{2T})\boldsymbol{v}^* + \tfrac{1}{2T}\boldsymbol{v}_1\big) - \lambda R(\boldsymbol{v}_1) \\
\leq & \sum_{t=1}^{T} - \ln((1 - \tfrac{1}{2T})(A\boldsymbol{v}^* + \boldsymbol{b})^\intercal \boldsymbol{x}_t) \\
& + \sum_{i=1}^{d} -\lambda \ln(\tfrac{1}{2T}(A\boldsymbol{v}_1 + \boldsymbol{b})^\intercal \boldsymbol{e}_i) - \lambda R(\boldsymbol{v}_1) \\
= & \sum_{t=1}^{T} f_t(\boldsymbol{v}^*) - T \ln(1 - \tfrac{1}{2T}) + d\lambda \ln(2T) \\
\leq & \sum_{t=1}^{T} f_t(\boldsymbol{v}^*) + \lambda d \ln(2T) + \frac{1}{2(1 - \frac{1}{2T})} \\
\leq & \sum_{t=1}^{T} f_t(\boldsymbol{v}^*) + \lambda d \ln(2T) + 1
\end{aligned}$$

Next, by using (2.16) of Nemirovski (2004) for the self-concordant function ϕ_t we find

$$\begin{aligned}
\phi_t(\boldsymbol{v}_t) - \phi_t(\boldsymbol{v}_{t+1}) &\leq - \ln(1 - \|\nabla f_t(\boldsymbol{v}_t)\|_t) - \|\nabla f_t(\boldsymbol{v}_t)\|_t \\
&\leq \|\nabla f_t(\boldsymbol{v}_t)\|_t^2,
\end{aligned}$$

where we used that $\nabla\phi_t(\boldsymbol{v}_t) = \nabla f_t(\boldsymbol{v}_t)$, $-\ln(1-s) - s \leq s^2$ for $s \in [0, \frac{2}{3}]$, and $\|\nabla f_t(\boldsymbol{v}_t)\|_t^2 \leq \frac{4}{9}$ by equation (7.2.1). To complete the proof we combine the above and find

$$\begin{aligned}
\mathcal{R}_T =& \phi_T(\boldsymbol{v}_{T+1}) - \lambda R(\boldsymbol{v}_1) - \sum_{t=1}^T f_t(\boldsymbol{v}^*) + \sum_{t=1}^T \big(\phi_t(\boldsymbol{v}_t) - \phi_t(\boldsymbol{v}_{t+1})\big) \\
\leq& \lambda d \ln(2T) + 1 + \sum_{t=1}^T \big(\phi_t(\boldsymbol{v}_t) - \phi_t(\boldsymbol{v}_{t+1})\big) \\
\leq& \lambda d \ln(2T) + 1 + \sum_{t=1}^T \|\nabla f_t(\boldsymbol{v}_t)\|_t^2.
\end{aligned}$$

□

Combining Lemma 30 with (7.2.1), we immediately see that the regret is bounded by

$$\mathcal{R}_T = O(\sqrt{dT \ln T}) \qquad \text{for } \lambda \approx \sqrt{\tfrac{T}{d \ln T}},$$

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 30.

7.2.2 Assuming Bounded Gradients

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

$$\begin{aligned}
\sum_{t=1}^T \|\nabla f_t(\boldsymbol{v}_t)\|_t^2 \leq& G^2 \sum_{t=1}^T \|\boldsymbol{y}_t\|_t^2 \\
\leq& G^2 \sum_{t=1}^T \boldsymbol{y}_t^\intercal \Big(\sum_{s=1}^t \boldsymbol{y}_s \boldsymbol{y}_s^\intercal + \lambda A^\intercal A\Big)^{-1} \boldsymbol{y}_t \\
=& O\Big(G^2 d \ln T\Big),
\end{aligned}$$

where the last step follows analogously to Hazan et al. (2007, Lemma 11) and using that $\det(A^\intercal A) = \det(\boldsymbol{I} + \boldsymbol{1}\boldsymbol{1}^\intercal) = (1 + \boldsymbol{1}^\intercal \boldsymbol{1}) \det(\boldsymbol{I}) = d$ by Sylvester's determinant theorem. This gives the optimal rate if G is small.

7.2.3 Source Coding and $\boldsymbol{x}_t$ in Finite Set

We call the case that $\boldsymbol{x}_t \in \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d\}$ the source coding setting. This case is easy to analyse, because $\boldsymbol{w}_t$ has a simple closed-form solution that coincides with Cover's universal portfolio algorithm. More generally, let us assume that $\boldsymbol{x}_t$ takes values in some finite set $\mathcal{X}$ of size k, so $k = d$ in the source coding setting, and let $n_t(\boldsymbol{x})$ denote the number of times that $\boldsymbol{x}_s = \boldsymbol{x}$ for $s \leq t$. Then

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

In particular, algorithm (7.1.1) achieves the optimal rate in the source coding setting.

7.2.4 A (Suboptimal) General Bound without Bounded Gradients

Since $R(\boldsymbol{v})$ 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 $\boldsymbol{w}_t$ by $m = O((\ln T)^d)$ sets $B_1, \ldots, B_m$ such that $\boldsymbol{w}^\intercal \boldsymbol{x} \leq 2\boldsymbol{u}^\intercal \boldsymbol{x}$ for all $\boldsymbol{w}, \boldsymbol{u} \in B_i$. It follows that

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

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

7.3 Discussion

The partial analysis presented above relies on (2.16) of Nemirovski (2004). An alternative approach could be to use (2.4) of Nemirovski (2004) instead, as is done by Bilodeau et al. (2020) to analyse the logarithmic loss. Another attempt could be made to improve the approach of Section 7.2.4 by employing other techniques from literature on self-concordant barriers. Inside the Dikin ellipsoid the hessians of self-concordant barriers are roughly proportional (see for example Proposition 2.3.2 by Nesterov and Nemirovskii (1994) or (2.2) by Nemirovski (2004)). Instead of covering the domain as described in Section 7.2.4 perhaps it is possible to cover the domain in Dikin ellipsoids more efficiently, although several unfruitful attempts have already been made.