Non-Convex Follow the Perturbed Leader
ECE 543 Final Project

Alan Yang

May 10, 2019

1 Introduction

Follow the Leader algorithms provide a natural extension of offline optimization for online learning. Recently, it was shown in [1] that the Follow the Perturbed Leader (FTPL) algorithm achieves regret that scales as $O(\sqrt{T})$ for Lipschitz, non-convex loss functions in the one-dimensional case. It was later shown in [4] that the $O(\sqrt{T})$ scaling holds in the general d-dimensional case as well. Interestingly, these regret bounds can also be shown for convex loss functions [3]. Regret bounds are derived in the 1-dimensional and general d-dimensional settings, with the assumption that the player has access to an α-accurate optimization oracle. It would seem that the convexity of the loss functions is not the crucial factor that makes online learning more difficult than offline statistical learning. In light of recent studies that interpret Generative Adversarial Network (GAN) training as an online learning problem [1], this is an encouraging result for good adversarial generative models.

1.1 Online Optimization

At each time t, the player selects an action f_t from $\mathcal{F} \subseteq \mathcal{H}$, where $\mathcal{H}$ is a d-dimensional Hilbert space and $\mathcal{F}$ has diameter D. The adversary then chooses an action z_t from $\mathcal{Z}$, and the player incurs loss $\ell_t(f_t) := \ell(f_t, z_t)$ where $\ell : \mathcal{F} \times \mathcal{Z} \to \mathbb{R}_+$. Assume that ℓ is L-Lipschitz in f with respect to the L1-norm, i.e. $\forall f_1, f_2 \in \mathcal{F}$ and any t, $|\ell_t(f_1) - \ell_t(f_2)| \leq L\|f_1 - f_2\|_1$.

The player aims to choose a sequence of actions (f_t) to minimize their regret, which is the cumulative loss with respect to the best fixed strategy $f^* = \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell_t(f)$:

$$R_T((f_t)) := \sum_{t=1}^{T} \ell_t(f_t) - \sum_{t=1}^{T} \ell_t(f^*) \tag{1}$$

It is assumed that the player is able to evaluate the loss functions and solve optimization problems of the form

$$\hat{f} = \arg\min_{f \in \mathcal{F}} \Big\{ \sum_{i=1}^{k} \ell_i(f) - \sigma^\top f \Big\} \tag{2}$$

with α-accuracy; that is, for some $\alpha \geq 0$ the resulting $\hat{f}$ satisfies

$$\sum_{i=1}^{k} \ell_i(\hat{f}) - \sigma^\top \hat{f} \leq \min_{f \in \mathcal{F}} \Big\{ \sum_{i=1}^{k} \ell_i(f) - \sigma^\top f \Big\} + \alpha \tag{3}$$

2 Follow the Leader Algorithms

Follow the Leader (FTL) algorithms extend offline optimization algorithms to the online case, and is used as a framework for understanding to what extent online problems are harder than static problems in statistical learning. At time t, the player chooses the action f_t that incurs the least historical risk, by solving optimization problems of the form

$$f_t = \arg\min_{f \in \mathcal{F}} \Big\{ \sum_{i=1}^{t-1} \ell_t(f) + \Omega(f) \Big\} \tag{4}$$

The "leader" is the best $f \in \mathcal{F}$ based on the player's observations thus far, and $\Omega(f)$ is a regularization term. In a sense, FTL is a natural extension of ERM to the online setting. As will be seen, the regularization term Ω is necessary to establish desirable regret bounds for FTL that are sublinear in T.

Lemma 1 (Be the Leader)**.** FTRL, as given in Equation 4, satisfies the following regret bound:

$$R_T((f_t)) \leq [\Omega(f^*) - \Omega(f_2)] + L \sum_{t=1}^{T} \|f_t - f_{t+1}\|_1 + \alpha T \tag{5}$$

The lemma and proof are similar to those of [3], modified for the α-accurate player, as in [4]. The proof is given in the Appendix.

3 Follow the Perturbed Leader

Assume that the adversary is oblivious: a fixed sequence $\ell_1, ...\ell_T$ is choisen beforehand, and the adversary cannot adapt to the player's choice of (f_t). This simplifies the analysis since the entries of σ can be drawn all at once at the beginning, and as a result Lemma 1 can be used to establish an expected regret bound. However, this restriction also does not change the main results: it was shown in [2] that regret bounds found in the oblivious case are equivalent asymptotically to those found in the general adversarial case.

Follow the Perturbed Leader (FTPL) introduces a stochastic regularization:

$$f_t = \arg\min_{f \in \mathcal{F}} \sum_{i=1}^{t-1} \ell_t(f) + \sigma^\top f \tag{6}$$

The noise is exponentially distributed i.i.d. with parameter η: $\sigma \sim \exp(\eta)^d$.

3.1 The 1-dimensional Case ($d = 1$)

Theorem 1 (Scalar Regret Bound)**.** Loss functions are L-Lipschitz, not necessarily convex. Consider the case $d = 1$. α-accurate FTPL satisfies the following expected regret bound:

$$\mathbb{E}[R_T((f_t))] \leq \frac{D}{\eta} + \frac{1000}{9} L^2 DT\eta + 2\alpha T \tag{7}$$

Moreover, for the choice $\eta = \sqrt{\frac{9}{1000L^2T}}$,

$$\mathbb{E}[R_T((f_t))] \leq \frac{2}{3}\sqrt{1000} DL\sqrt{T} + 2\alpha T \tag{8}$$

This result shows that in the 1-d case, FTPL achieves expected regret that scales as $O(\sqrt{T})$, which is as good as the rate in the convex case. Furthermore, Equation 7 intuitively demonstrates the role of the additive noise. Smaller values of η correspond to higher noise levels, since the noise is exponentially distributed with parameter η. Increasing noise stabilizes the player's actions, but with an accuracy penalty. The notation $f_t(\sigma)$ is used to emphasize that $f_t(\sigma)$ was chosen for a particular instantiation of σ. We first begin with a few preliminaries.

Definition 1 (One-step differences)**.** For each t and each dimension i, denote the differences:

$$f_{min,t}^{(i)}(\sigma) = \min\{f_t^{(i)}(\sigma), f_{t+1}^{(i)}(\sigma)\} \tag{9}$$

$$f_{max,t}^{(i)}(\sigma) = \max\{f_t^{(i)}(\sigma), f_{t+1}^{(i)}(\sigma)\} \tag{10}$$

The superscripts are dropped in the 1-d case, since the dimension is clear in context.

Lemma 2 (Monotonicity Property [4])**.** Consider α-accurate FTPL. Denote the i-th standard basis vector as $\boldsymbol{e}_i$. At time t, the player chooses $f_t(\sigma)$. Then for any $c \geq 0$,

$$f_t^{(i)}(\sigma + c\boldsymbol{e}_i) \geq f_t^{(i)}(\sigma) - \frac{2\alpha}{c} \tag{11}$$

Lemma 3 (Stability w.r.t. σ [4])**.** Denote the i-th standard basis vector as $\boldsymbol{e}_i$. Suppose that $\|f_t(\sigma) - f_{t+1}(\sigma)\|_1 \leq 10d|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)|$. Then for a choice of $\sigma' = \sigma + 100Ld\boldsymbol{e}_i$,

$$f_{min,t}^{(i)}(\sigma') \geq f_{max,t}^{(i)}(\sigma) - \frac{1}{10}|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)| - \frac{3\alpha}{100Ld} \tag{12}$$

The purpose of Definition 1 is to establish how quickly the player's choices are able to change between time t and $t+1$, for a given noise level. Lemma 2 relates changes in noise level to changes in the player's action. It is not directly used in the proof for the 1-d case, but is used to prove Lemma 3 as well as Theorem 2. The proof for both the 1-d and general d-dimensional case relies on Lemma 3. This lemma is used to establish a bound on the expected stability of the player's actions between times, in the style of Lemma 1. It does so by characterizing how the player's actions change in response to change in noise, from σ to σ'.

Note that Lemma 3 and its proof require a good choice of σ' relative to σ. As presented, it uses the choices presented in [4], which gives a version of the lemma that is able to prove $O(\sqrt{T})$ bounds on the regret. This choice of σ, σ' are used to show a regret bound for the 1-d case similar to the bound presented in [1], except extended to the case where the player's optimization is only α-accurate, and with the version of Lemma 3 presented here.

Also, the version of Lemma 3 that appears in [1] does not contain the condition on $\|f_t(\sigma) - f_{t+1}(\sigma)\|_1$. Including this condition changes the structure of the proof between the 1-dimensional and d-dimensional versions of the regret bound, but enables an $O(\sqrt{T})$ bound on the regret, which was not achievable in [1]. Proofs of Lemmas 2 and 3 use α-optimality condition given in Equation 3 as well as the fact that the loss functions are L-Lipschitz.

Proof of Theorem 1. For an oblivious adversary, the expected regret of FTPL can be bounded using Lemma 1:

$$\mathbb{E}[R_T((f_t))] \leq \mathbb{E}[\sigma(f_1 - f^*)] + L \sum_{t=1}^{T} \mathbb{E}\big[|f_t - f_{t+1}|\big] + \alpha T \tag{13}$$

$$\leq \frac{D}{\eta} + L \sum_{t=1}^{T} \mathbb{E}\big[|f_t - f_{t+1}|\big] + \alpha T \tag{14}$$

The second inequality is true since σ is exponentially distributed and D is the diameter of $\mathcal{F}$. The term $\mathbb{E}\big[|f_t - f_{t+1}|\big]$ can now be bounded using Lemmas 2 and 3. Fix a noise value σ and $\sigma' = \sigma + 100L$. Now we have

$$\mathbb{E}\big[|f_t(\sigma) - f_{t+1}(\sigma)|\big] = \mathbb{E}[f_{max,t}(\sigma)] - \mathbb{E}[f_{min,t}(\sigma)] \tag{15}$$

One way to bound the right hand side is to find a lower bound of $\mathbb{E}[f_{min,t}(\sigma)]$ using $\mathbb{E}[f_{max,t}(\sigma)]$.

$$\begin{aligned}
\mathbb{E}[f_{min,t}] &= \mathbb{E}[f_{min,t}(\sigma)|\sigma < 100L]P[\sigma < 100L] + \mathbb{E}[f_{min,t}(\sigma)|\sigma \geq 100L]P[\sigma \geq 100L] \\
&= \mathbb{E}[f_{min,t}(\sigma)|\sigma < 100L](1 - e^{-\eta 100L}) + \mathbb{E}[f_{min,t}(\sigma')|\sigma' \geq 0]e^{-\eta 100L} \\
&\overset{(a)}{\geq} (\mathbb{E}[f_{max,t}(\sigma)] - D)(1 - e^{-\eta 100L}) + \mathbb{E}[f_{min,t}(\sigma')]e^{-\eta 100L} \\
&\overset{(b)}{\geq} (\mathbb{E}[f_{max,t}(\sigma)] - D)(1 - e^{-\eta 100L}) + \Big[\mathbb{E}[f_{max,t}(\sigma)] - \frac{\mathbb{E}\big[|f_t(\sigma) - f_{t+1}(\sigma)|\big]}{10} - \frac{3\alpha}{100L}\Big]e^{-\eta 100L} \\
&= \mathbb{E}[f_{max,t}(\sigma)] - D(1 - e^{-\eta 100L}) - \Big[\frac{\mathbb{E}\big[|f_t(\sigma) - f_{t+1}(\sigma)|\big]}{10} + \frac{3\alpha}{100L}\Big]e^{-\eta 100L} \\
&\overset{(c)}{\geq} \mathbb{E}[f_{max,t}(\sigma)] - \eta 100LD - \Big[\frac{\mathbb{E}\big[|f_t(\sigma) - f_{t+1}(\sigma)|\big]}{10} + \frac{3\alpha}{100L}\Big]
\end{aligned}$$

(a) uses the fact that $\mathcal{F}$ has diameter D, (b) uses Lemma 3, and (c) the fact that $e^x \geq 1 + x$. Note that in the 1-d case, the condition in Lemma 3 holds trivially. Combining with Equation 15 gives

$$\mathbb{E}\big[|f_t(\sigma) - f_{t+1}(\sigma)|\big] \leq \frac{1000}{9} LD\eta + \frac{\alpha}{30L} \tag{16}$$

Plugging into Equation 14 gives Equation 7 as desired, and optimizing over η gives Equation 8. $\square$

3.2 The General d-dimensional Case

The proof for the 1-dimensional case can be extended to the d-dimensional case by starting with the fact that $\mathbb{E}[\|f_t - f_{t+1}\|_1] = \sum_{k=1}^{d} \mathbb{E}[|f_t - f_{t+1}|]$ and finding a bound for each dimension each separately. Similar to the 1-d case, the idea is to use Lemma 3 to bound this term. Lemma 4 bounds the regularization terms, and is proved in the Appendix.

Theorem 2 (Non-convex FTPL Regret Bound)**.** Loss functions are L-Lipschitz, not necessarily convex. Consider the case $d = 1$. α-accurate FTPL satisfies the following expected regret bound:

$$\mathbb{E}[R_T((f_t))] \leq \frac{D(\log d + 1)}{\eta} + 125L^2d^2DT\eta + 2\alpha T \tag{17}$$

Moreover, for the choice $\eta = \sqrt{\frac{\log d+1}{125L^2d^2T}}$,

$$\mathbb{E}[R_T((f_t))] \leq 2\sqrt{125(\log d + 1)}LdD\sqrt{T} + 2\alpha T \tag{18}$$

Lemma 4. Fix $f_1, f_2 \in \mathcal{F}$, where $\mathcal{F}$ has diameter D, and suppose that $\sigma \sim \exp(\eta)^d$. Then

$$\mathbb{E}[\sigma^\top(f_1 - f_2)] \leq \frac{D(\log(d) + 1)}{\eta} \tag{19}$$

Proof of Theorem 2. The proof proceeds almost identically to that for the 1-d case, except that a couple of modifications are added in order to make use of Lemma 3. Combining Lemmas 1 and 4 gives

$$\mathbb{E}[R_T((f_t))] = \frac{D(\log(d) + 1)}{\eta} + L\sum_{t=1}^{T}\mathbb{E}\big[\|f_t - f_{t+1}\|_1\big] + \alpha T \tag{20}$$

The idea is to start by considering each dimension of σ separately, holding the others fixed. To this end, define for each dimension $i = 1, ..., d$:

$$\mathbb{E}_{\setminus i}[|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)|] := \mathbb{E}\Big[|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)|\Big|\{\sigma_j\}_{j\neq i}\Big] \tag{21}$$

If a bound can be found for this conditional expectation that holds for each dimension i, it will also hold for the unconditioned expectation. Considering each dimension separately allows for a reduction to the 1-d case. In order to make use of Lemma 3, define the following event, as well as its associated conditional probability:

$$\mathcal{E} = \{\sigma : \|f_t(\sigma) - f_{t+1}(\sigma)\|_1 \leq 10d|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)|\} \tag{22}$$

$$P_{\setminus i}[\mathcal{E}] = P[\mathcal{E}|\{\sigma_j\}_{j\neq i}] \tag{23}$$

Analogous to the 1-d case, the goal is to find a lower bound on $\mathbb{E}_{\setminus i}[f_{min,t}(\sigma)]$ using $\mathbb{E}_{\setminus i}[f_{max,t}(\sigma)]$, since applying Equation 15 gives:

$$\mathbb{E}_{\setminus i}\big[|f_t^{(i)}(\sigma) - f_{t+1}^{(i)}(\sigma)|\big] = \mathbb{E}_{\setminus i}[f_{max,t}^{(i)}(\sigma)] - \mathbb{E}_{\setminus i}[f_{min,t}^{(i)}(\sigma)] \tag{24}$$

Fix σ and $\sigma' = \sigma + 100Ld\boldsymbol{e}_i$.

$$
\begin{aligned}
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)] &= \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)|\sigma_i < 100Ld]P[\sigma_i < 100Ld] + \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)|\sigma_i \geq 100Ld]P[\sigma_i \geq 100Ld] \\
&= \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)|\sigma_i < 100Ld](1 - e^{-\eta 100Ld}) + \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')]e^{-\eta 100Ld} \\
&\geq (\mathbb{E}_{\backslash i}[f^{(i)}_{max,t}(\sigma)] - D)(1 - e^{-\eta 100Ld}) + \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')]e^{-\eta 100Ld} \qquad (25)
\end{aligned}
$$

The inequality uses the fact that $\mathcal{F}$ has diameter D, as was done in the 1-d case. The second conditional expectation can be expanded:

$$
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')] = \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')|\mathcal{E}]P_{\backslash i}[\mathcal{E}] + \mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')|\mathcal{E}^c]P_{\backslash i}[\mathcal{E}^c] \qquad (26)
$$

The individual expectations can be bounded using Lemmas 3 and 2 respectively, as well as the definition of $\mathcal{E}^c$:

$$
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')|\mathcal{E}] \geq \mathbb{E}_{\backslash i}\left[f^{(i)}_{max,t}(\sigma) - \frac{1}{10}|f^{(i)}_t(\sigma) - f^{(i)}_{t+1}(\sigma)| - \frac{3\alpha}{100L}\middle|\mathcal{E}\right] \qquad (27)
$$

$$
\begin{aligned}
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma')|\mathcal{E}^c] &\geq \mathbb{E}_{\backslash i}\left[f^{(i)}_{min,t}(\sigma) - \frac{2\alpha}{100Ld}\middle|\mathcal{E}^c\right] \\
&\geq \mathbb{E}_{\backslash i}\left[f^{(i)}_{max,t}(\sigma) - \frac{1}{10d}\|f_t(\sigma) - f_{t+1}(\sigma)\|_1 - \frac{2\alpha}{100Ld}\middle|\mathcal{E}^c\right] \qquad (28)
\end{aligned}
$$

Combining equations 25, 26, 27, and 28, and noting that the dependencies on $\mathcal{E}$ and $\mathcal{E}^c$ cancel gives:

$$
\begin{aligned}
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)] \geq\ & (\mathbb{E}_{\backslash i}[f^{(i)}_{max,t}(\sigma)] - D)(1 - e^{-\eta 100Ld}) \\
&+ e^{-\eta 100Ld}\Bigg(\mathbb{E}_{\backslash i}\left[f^{(i)}_{max,t}(\sigma) - \frac{3\alpha}{100Ld}\right] \\
&- \mathbb{E}_{\backslash i}\left[\frac{1}{10}\left|f^{(i)}_t(\sigma) - f^{(i)}_{t+1}(\sigma)\right| + \frac{1}{10d}\|f_t(\sigma) - f_{t+1}(\sigma)\|_1\right]\Bigg) \qquad (29)
\end{aligned}
$$

Again using the fact that $e^x \geq 1 + x$,

$$
\begin{aligned}
\mathbb{E}_{\backslash i}[f^{(i)}_{min,t}(\sigma)] \geq\ & \mathbb{E}_{\backslash i}[f^{(i)}_{max,t}(\sigma)] - \eta 100LdD - \frac{3\alpha}{100Ld} \\
&- \mathbb{E}_{\backslash i}\left[\frac{1}{10}\left|f^{(i)}_t(\sigma) - f^{(i)}_{t+1}(\sigma)\right| + \frac{1}{10d}\|f_t(\sigma) - f_{t+1}(\sigma)\|_1\right] \qquad (30)
\end{aligned}
$$

Combining with Equation 21 and rearranging gives a result similar to in the 1-d case,

$$
\mathbb{E}_{\backslash i}\left|f^{(i)}_t(\sigma) - f^{(i)}_{t+1}(\sigma)\right| \leq \frac{1000}{9}\eta LdD + \frac{1}{9d}\mathbb{E}_{\backslash i}\|f_t(\sigma) - f_{t+1}(\sigma)\|_1 + \frac{\alpha}{30Ld} \qquad (31)
$$

Since this holds for every i with $\sigma_j, j \neq i$ fixed, this bound also holds for the unconditioned expectation. Summing over the d dimensions gives

$$
\mathbb{E}\left\|f^{(i)}_t(\sigma) - f^{(i)}_{t+1}(\sigma)\right\|_1 \leq 125\eta Ld^2 D + \frac{\alpha}{20L} \qquad (32)
$$

Plugging this into Equation 20 gives Equation 17 and optimizing over η gives Equation 18. □

References

[1] Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization. *ArXiv e-prints*, March 2019. arXiv:1810.07362v6 [cs.LG].

[2] Nicolo Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, USA, 2006.

[3] Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *J. Comput. Syst. Sci.*, 71(3):291–307, October 2005.

[4] Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. *ArXiv e-prints*, March 2019. arXiv:1903.08110v1 [cs.LG].

4 Appendix

4.1 Proof of Lemma 1

The regret of playing the sequence (f_t) is

$$\begin{aligned} R_T((f_t)) &= \sum_{t=1}^{T}[\ell_t(f_t) - \ell_t(f^*)] \\ &= \sum_{t=1}^{T}[\ell_t(f_t) - \ell_t(f_{t+1})] + \sum_{t=1}^{T}[\ell_t(f_{t+1}) - \ell_t(f^*)] \\ &\leq L\sum_{t=1}^{T}\|f_t - f_{t+1}\|_1 + \sum_{t=1}^{T}[\ell_t(f_{t+1}) - \ell_t(f^*)] \\ &\leq L\sum_{t=1}^{T}\|f_t - f_{t+1}\|_1 + [\Omega(f^*) - \Omega(f_2)] + \alpha T \end{aligned}$$

The first inequality follows from the fact that ℓ is L-Lipschitz, and the second can be shown by induction. Consider the base case $T = 1$. Since f_2 is an α-accurate minimizer, for any f^*

$$\begin{aligned} \ell_1(f_2) - \Omega(f_2) &\leq \inf_f[\ell_1(f) - \Omega(f)] + \alpha \\ &\leq \ell_1(f^*) - \Omega(f^*) + \alpha \end{aligned}$$

and therefore $\ell_1(f_2) - \ell_1(f^*) \leq [\Omega(f_2) - \Omega(f^*)] + \alpha$.

Next, suppose $\sum_{t=1}^{T}[\ell_t(f_{t+1}) - \ell_t(f^*)] \leq [\Omega(f^*) - \Omega(f_2)] + \alpha T$ for any f^* and $T \leq T_0 - 1$.

$$\begin{aligned}
\sum_{t=1}^{T_0-1} \ell_t(f_{t+1}) &\overset{(a)}{\leq} \Big[\sum_{t=1}^{T_0-1} \ell_t(f_{T_0+1}) + [\Omega(f_{T_0+1}) - \Omega(f_2)] + \alpha(T_0 - 1)\Big] + \ell_{T_0}(f_{T_0+1}) \\
&= \Big[\sum_{t=1}^{T_0} \ell_t(f_{T_0+1}) - \Omega(f_{T_0+1})\Big] + \Omega(f_2) + \alpha(T_0 - 1) \\
&\overset{(b)}{\leq} \sum_{t=1}^{T_0} \ell_t(f_{T_0+1}) + [\Omega(f_2) - \Omega(f_{T_0+1})] + \alpha T_0 \\
&= \sum_{t=1}^{T_0} \ell_t(f^*) + [\Omega(f^*) - \Omega(f_2)] + \alpha T_0
\end{aligned}$$

(a) uses the assumption, and (b) uses the fact that f_{T_0+1} is an α-approximate minimizer of R_{T_0}.

4.2 Proof of Lemma 4

$$\begin{aligned}
\mathbb{E}[\sigma^\top(f_1 - f_2)] &\leq \mathbb{E}[\|\sigma^\top(f_1 - f_2)\|_1] \\
&\overset{(a)}{\leq} \mathbb{E}[\|\sigma\|_\infty \cdot \|f_1 - f_2\|_1] \\
&\overset{(b)}{\leq} \mathbb{E}[\|\sigma\|_\infty] \cdot D \\
&\overset{(c)}{\leq} \frac{D(\log(d) + 1)}{\eta}
\end{aligned}$$

(a) follows from Hölder's Inequality, and (b) from the restriction on $\mathcal{F}$. (c) uses the fact that $\mathbb{E}\|\sigma\|_\infty \leq \frac{D(\log d+1}{\eta}$, for exponential σ [1].