FTL FOR OQO PROBLEMS

- One popular instantiation of the online learning problem is the problem of online quadratic optimization (OQO).
- In its most general form, the loss function is thereby defined as

$$L(a_1, z_1) = \frac{1}{2} ||a_1 - z_1||_{2}^{2},$$

where $A, Z \subset \mathbb{R}^d$.

Proposition: Using FTL on any online quadratic optimization problem with A = ℝ^d and V = sup ||z||_{z∈Z}, leads to a regret of

$$R_T^{\text{PTL}} \le 4V^2((\log(T) + 11)).$$



Proof:

In the following, we denote a₁^{FTL}, a₂^{FTL}, ...si simply by₁, a₂, ...
 a₁, a₂, ...



Proof:

In the following, we denote a₁^{FTL}, a₂^{FTL}, ...si simply by₁, a₂, ...

```
\begin{array}{c} \textbf{a_{1}, a_{2}, \dots} \\ & \textbf{Reminder (Useful Lemma):} \\ & \textbf{Reminder (Useful Lemma):} \quad T \\ & R_{T}^{\text{FIL}} \leq \sum_{t=1}^{T_{1}} L(a_{t}^{\text{FIL}}, z_{t}) - \sum_{t=1}^{T_{1}} L(a_{t+1}^{\text{FIL}}, z_{t}) \\ & R_{T}^{\text{FIL}} \leq \sum_{t=1}^{T_{1}} (a_{t}^{\text{FIL}}, z_{t}) - \sum_{t=1}^{T_{1}} (a_{t+1}^{\text{FIL}}, z_{t}) \end{array}
```



Proof:

In the following, we denote a₁^{FTL}, a₂^{FTL}, ...si simply by₁, a₂, ...

```
Reminder (U seful Lemma):

Reminder (U seful Lemma):

H_{T}^{\text{FTL}} \leq \sum_{l=1}^{T} \frac{L(a_{l+1}^{\text{FTL}}, z_{l})}{L(a_{l+1}^{\text{FTL}}, z_{l})}

Using this Remark of the North Path (a_{l+1}^{\text{FTL}}, z_{l})
```

Using this lemma, we just have to show that

$$\sum_{t=1}^{T} (L(a_t, z_t) - L(a_{t+1}, z_t)) \le 4L^2 \cdot (\log(T) + 1).$$
 (1)
$$\sum_{t=1}^{T} ((a_t, z_t) - (a_{t+1}, z_t)) \le 4L^2 \cdot (\log(T) + 1).$$
 (1)



Proof:

In the following, we denote a₁^{FTIL}, a₂^{FTIL}, ...si simply by₁, a₂, ...

```
Reminder (U seful Lemma):

Reminder (U seful Lemma):

R_{T}^{\text{Eminder}} \leq L(a_{t-1}^{\text{FTL}}, z_{t}) - \sum_{L=1}^{L} L(a_{t-1}^{\text{FTL}}, z_{t})

Using this Remma, Lequit (Reve to both pate (a_{t+1}^{\text{FTL}}, z_{t}))
```

Using this lemma, we just have to show that

$$\sum_{t=1}^{T} (L(a_t, z_t) - L(a_{t+1}, z_t)) \le 4L^2 \cdot (\log(T) + 1) . \tag{1}$$

$$\sum_{t=1}^{T} ((a_t, z_t) - (a_{t+1}, z_t)) \le 4L^2 \cdot (\log(T) + 1) . \tag{1}$$

- So, we will prove (1). For this purpose, we compute the
- So, we will prove (1) For this purpose, we compute the
 explicit form of the actions of FTL for this type of online
 learning problem.



• Claim: It holds that $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$, if $(a_s z) = \frac{1}{2} ||a_s - z||_{2}^{2/2}$.



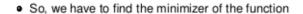
- Claim: It holds that $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$, if $(a_t z) = \frac{1}{2} ||a_t z||_{2}^{2/2}$.
 - Recall that

$$a_t^{\mathrm{FTL}} a_t^{\mathrm{FTL}} = \underset{a \in \mathcal{A}}{\operatorname{main}} \min \sum_{s = 1}^{t-1} \mathcal{L}(a, z_s) = \underset{a \in \mathcal{A}}{\operatorname{argmin}} \sum_{s = 1}^{t-1} \sum_{s = 1}^{t-1} a_2^1 |z_s||_2^2 z_s||_2^2 \,.$$



- Claim: It holds that $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$, if $(a_s z) = \frac{1}{2} ||a_s z||_{22}^{22}$
 - Recall that

$$a_t^{\mathrm{FTL}} a_t^{\mathrm{FTL}} \underset{a \in \mathcal{A}}{\operatorname{drg-nairg.min}} \sum_{ss = 1}^{t-1} l(a, z_s) = \underset{a \in \mathcal{A}}{\operatorname{arg.min}} \sum_{s = 1}^{t-1} \underbrace{2^{t-1}_{s-1}}_{s-1} a_2^1 ||z_s||_2^2 z_s||_2^2 \,.$$



$$f(a) := \sum_{s=1}^{t-1} \frac{1}{2} \|a - z_s\|_2^2 = \sum_{s=1}^{t-1} \frac{1}{2} (a - z_s)^{\top} (a - z_s).$$



- Claim: It holds that $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$, if $(a_t z_t) = \frac{1}{2} ||a_t z||_{22}^{22}$
 - Recall that

$$a_t^{\mathrm{FTL}} a_t^{\mathrm{FTL}} \underset{a \in \mathcal{A}}{\operatorname{arg-nairg}} \min \sum_{s s = 1}^{t-i-1} \mathcal{L}(a, z_s) = \underset{a \in \mathcal{A}}{\operatorname{arg-min}} \sum_{s = 1}^{t-1} \underbrace{2^{t-1}_{s-1}}_{s-1} a_2^{t} ||z_s||_2^2 z_s||_2^2.$$

. So, we have to find the minimizer of the function

$$f(a) := \sum_{s=1}^{t-1} \frac{1}{2} \|a - z_s\|_2^2 = \sum_{s=1}^{t-1} \frac{1}{2} (a - z_s)^{\top} (a - z_s).$$

• Compute $\nabla f(a) = \sum_{s=1}^{t-1} a - z_s = (t-1)a - \sum_{s=1}^{t-1} z_s$, which we set to zero and solve with respect to a to obtain the claim.

(f is convex, so that this leads indeed to a minimizer.)



 Hence, a_t is the empirical average of z₁,..., z_{t-1} and we can provide the following incremental update formula for its computation

$$a_{t+1} = \frac{1}{t} \cdot \sum_{s=1}^{t} z_s = \frac{1}{t} \left(z_t + \sum_{s=1}^{t-1} z_s \right)$$
$$= \frac{1}{t} (z_t + (t-1)a_t) = \frac{1}{t} z_t + (1 - \frac{1}{t}) a_t.$$



$$a_{l+1} - z_l = (1 - \frac{1}{l}) \cdot a_l + \frac{1}{l} z_l - z_l = (1 - \frac{1}{l}) \cdot (a_l - z_l).$$

Claim:

$$L(a_t, z_t) - (a_{t+1}, z_t) \leq \frac{1}{t} ||a_t - z_t||_{2}^{2/2}.$$
 (2)



Reminder:
$$a_{t+1} - z_t = \left(1 - \frac{1}{t}\right) \cdot \left(a_t - z_t\right)$$
.

Indeed, this can be seen as follows

$$\begin{split} L(a_{r},z_{r}) - \left(\hat{a}_{R+1},z_{r}\right) &= \frac{1}{2} \left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} \frac{1}{2} - \frac{11}{22} \left\| \hat{a}_{R+1} - z_{r} \right\|_{2}^{2} \frac{1}{2} \\ &= \frac{1}{2} \left(\left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} \frac{1}{2} - \left\| \hat{a}_{R+1} - z_{r} \right\|_{2}^{2} \right) \\ &= \frac{1}{2} \left(\left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} \frac{1}{2} - \left\| \left(1 - \frac{1}{r} \right) \cdot \left(\hat{a}_{R} - z_{r} \right) \right\|_{2}^{2} \right) \right). \end{split}$$



Reminder:
$$a_{t+1} - z_t = \left(1 - \frac{1}{t}\right) \cdot \left(a_t - z_t\right)$$
.

Indeed, this can be seen as follows

$$\begin{split} L(a_{r},z_{r}) - \left(\hat{a}_{R+1},z_{r}\right) &= \frac{1}{2} \left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} - \frac{11}{2^{2}} \left\| \hat{a}_{R+1} - z_{r} \right\|_{2}^{2} \right\} \\ &= \frac{1}{2} \left(\left(\left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} - \left\| \left| \hat{a}_{R+1} - z_{r} \right|_{2}^{2} \right) \right) \\ &= \frac{1}{2} \left(\left(\left\| \hat{a}_{R} - z_{r} \right\|_{2}^{2} - \left\| \left(\left(\left(\left| \frac{1}{r} \right| \right) \right) \left(\left| \hat{a}_{R} - z_{r} \right|\right)_{2}^{2} \right) \right). \end{split}$$

· And from this,

$$\begin{split} L(\mathbf{a}_{1}, z_{1}) - (\mathbf{a}_{1}, z_{1}) &= \frac{1}{2} \left[\left(||\mathbf{a}_{1} - z_{1}||_{2}^{2} - \left(1 - \frac{1}{t} \right)^{2} \right)^{2} ||\mathbf{a}_{1} - z_{1}||_{2}^{2} \right) \\ &= \frac{1}{2} \left[\left(1 - \left(1 - \frac{1}{t} \right)^{2} \right) \right) ||\mathbf{a}_{1} - z_{1}||_{2}^{2} \\ &= \left(\frac{1}{t} \frac{1}{t} - \frac{1}{2t} \right) \right) ||\mathbf{a}_{2} - z_{1}||_{2}^{2} \\ &\leq \frac{1}{t} \frac{1}{t} ||\mathbf{a}_{2} - z_{1}||_{2}^{2} \end{split}$$



RReminder:
$$L(a_t, z_t) - (a_{t+1}, z_t) \leq \frac{1}{t} ||a_{t+1} - z_t||_{2}^{2}$$
. (2)

Since by assumption L = sup_{z∈Z} ||z||₂ and a_t is the empirical average of z₁,..., z_{t-1}, we have that ||a_t||₂ ≤ L.



FReminder:
$$L(a_t, z_t) - (a_{t+1}, z_t) \leq \frac{1}{t} ||a_t - z_t||_{2}^{2}$$
 (2)

- Since by assumption L = sup_{z∈Z} ||z||₂ and a_t is the empirical average of z₁,..., z_{t-1}, we have that ||a_t||₂ ≤ L.
- ullet Now the triangle inequality states that for any two vectors $x,y\in\mathbb{R}^d$ it holds that

$$||x + y||_2 \le ||x||_2 + ||y||_2$$

so that

$$||a_t - z_t||_2 \le ||a_t||_2 + ||z_t||_2 \le 2L.$$
 (3)



FReminder:
$$L(a_t, z_t) - (a_{t+1}, z_t) \leq \frac{1}{t} \frac{1}{t} ||a_{t} - z_t||_{2}^{2}$$
. (2)

- Since by assumption L = sup_{z∈Z} ||z||₂ and a_t is the empirical average of z₁,..., z_{t-1}, we have that ||a_t||₂ ≤ L.
- ullet Now the triangle inequality states that for any two vectors $x,y\in\mathbb{R}^d$ it holds that

$$||x + y||_2 \le ||x||_2 + ||y||_2$$

so that

$$||a_t - z_t||_2 \le ||a_t||_2 + ||z_t||_2 \le 2L.$$
 (3)

Summing over all t in (2) and using (3) we arrive at

$$\begin{split} \sum_{t=t+1}^{TT} \left(\left(a_t, z_t \right) - \left(a_{t+1}, z_t \right) \right) &\leq \sum_{t=t+1}^{TT} \left(\frac{1}{t} \frac{1}{t} \left\| \left| a_{t} - z_t \right|_2^2 \right) \right) \leq \sum_{t=t+1}^{TT} \frac{1}{t} \left(2L \right)^2 \\ &= 4L^2 \cdot \sum_{t=t+1}^{TT} \frac{1}{t} \cdot \left(2L \right)^2 \end{split}$$



FReminder:
$$\sum_{t=t+1}^{TT} ((a_t, z_t) - (a_{t+1}, z_t)) \leq 4L^{2/2} \sum_{t=t+1}^{TT} \frac{1}{t}$$



FReminder:
$$\sum_{t=t+1}^{TT} ((a_t, z_t) - (a_{t+1}, z_t))) \leq 4L^{22} \sum_{t=1}^{TT} \frac{1}{t}$$



• Now, it holds that $\sum_{t=1}^{T} \frac{1}{t} \leq \log(T) + 1$, so that we obtain

$$\sum_{t=1}^{T} ((a_{t}, z_{t})) - (a_{t}(a_{t}, z_{t})) \leq 4L^{2}L \sum_{t=1}^{T} \stackrel{?}{\leq} 4L^{2}L^{2}(\log(T) + 1),$$

which is what we wanted to prove.