# 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

$$(a_t, z_t) = \frac{1}{2} ||a_t - z_t||_2^2$$

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

• **Proposition:** Using FTL on any online quadratic optimization problem with  $\mathcal{A} = \mathbb{R}^d$  and  $V = \sup_{z \in \mathcal{Z}} ||z||_2$ , leads to a regret of

$$R_T^{\rm FTL} \leq 4V^2 \left(\log(T) + 1\right).$$

× 0 0 × × ×

- Proof:
  - In the following, we denote  $a_1^{\text{FTL}}, a_2^{\text{FTL}}, \dots$  simply by  $a_1, a_2, \dots$

× 0 0 × 0 × ×

#### • Proof:

• In the following, we denote  $a_1^{\text{FTL}}, a_2^{\text{FTL}}, \dots$  simply by  $a_1, a_2, \dots$ 

Reminder (Useful Lemma):

$$R_T^{\text{FTL}} \leq \sum_{t=1}^T \left( a_t^{\text{FTL}}, z_t \right) - \sum_{t=1}^T \left( a_{t+1}^{\text{FTL}}, z_t \right)$$

× 0 0 × 0 × ×

#### • Proof:

• In the following, we denote  $a_1^{\text{FTL}}, a_2^{\text{FTL}}, \dots$  simply by  $a_1, a_2, \dots$ 

Reminder (Useful Lemma):  $R_T^{\text{FTL}} \leq \sum_{t=1}^T (a_t^{\text{FTL}}, z_t) - \sum_{t=1}^T (a_{t+1}^{\text{FTL}}, z_t)$ 

• Using this lemma, we just have to show that

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



#### • Proof:

• In the following, we denote  $a_1^{\text{FTL}}, a_2^{\text{FTL}}, \dots$  simply by  $a_1, a_2, \dots$ 

Reminder (Useful Lemma):

$$R_T^{\text{FTL}} \leq \sum_{t=1}^{\prime} \left( a_t^{\text{FTL}}, z_t \right) - \sum_{t=1}^{\prime} \left( a_{t+1}^{\text{FTL}}, z_t \right)$$

• Using this lemma, we just have to show that

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

• 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  $(a, z) = \frac{1}{2} ||a - z||_2^2$ .

× 0 0 × 0 × ×

• Claim: It holds that  $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$ , if  $(a, z) = \frac{1}{2} ||a - z||_2^2$ . • Recall that

$$a_t^{\text{FTL}} = \operatorname*{arg\,min}_{a \in \mathcal{A}} \sum_{s=1}^{t-1} (a, z_s) = \operatorname*{arg\,min}_{a \in \mathcal{A}} \sum_{s=1}^{t-1} \frac{1}{2} \left\| a - z_s \right\|_2^2.$$

× 0 0 × × ×

• Claim: It holds that  $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$ , if  $(a, z) = \frac{1}{2} ||a - z||_2^2$ . • Recall that

$$a_t^{\text{FTL}} = \arg\min_{a \in \mathcal{A}} \sum_{s=1}^{t-1} (a, z_s) = \arg\min_{a \in \mathcal{A}} \sum_{s=1}^{t-1} \frac{1}{2} ||a - z_s||_2^2.$$

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

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

X

 $\times \times$ 

• Claim: It holds that  $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$ , if  $(a, z) = \frac{1}{2} ||a - z||_2^2$ . • Recall that

$$a_t^{\text{FTL}} = \arg\min_{a \in \mathcal{A}} \sum_{s=1}^{t-1} (a, z_s) = \arg\min_{a \in \mathcal{A}} \sum_{s=1}^{t-1} \frac{1}{2} ||a - z_s||_2^2.$$

• 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.)

0 0 X X 0 X X

• Hence, *a<sub>t</sub>* is the empirical average of *z*<sub>1</sub>,..., *z*<sub>*t*-1</sub> and we can provide the following incremental update formula for its computation

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

× × ×

• Hence, *a<sub>t</sub>* is the empirical average of *z*<sub>1</sub>,..., *z*<sub>*t*-1</sub> and we can provide the following incremental update formula for its computation

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

From the last display we derive that

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

• Hence, *a<sub>t</sub>* is the empirical average of *z*<sub>1</sub>,..., *z*<sub>*t*-1</sub> and we can provide the following incremental update formula for its computation

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

• From the last display we derive that

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

• Claim:

$$(a_t, z_t) - (a_{t+1}, z_t) \leq \frac{1}{t} \cdot ||a_t - z_t||_2^2.$$
 (2)

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

Indeed, this can be seen as follows

$$\begin{aligned} (a_t, z_t) - (a_{t+1}, z_t) &= \frac{1}{2} ||a_t - z_t||_2^2 - \frac{1}{2} ||a_{t+1} - z_t||_2^2 \\ &= \frac{1}{2} \left( ||a_t - z_t||_2^2 - ||a_{t+1} - z_t||_2^2 \right) \\ &= \frac{1}{2} \left( ||a_t - z_t||_2^2 - ||(1 - \frac{1}{t}) \cdot (a_t - z_t)||_2^2 \right). \end{aligned}$$



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

Indeed, this can be seen as follows

$$(a_t, z_t) - (a_{t+1}, z_t) = \frac{1}{2} ||a_t - z_t||_2^2 - \frac{1}{2} ||a_{t+1} - z_t||_2^2$$
  
=  $\frac{1}{2} \left( ||a_t - z_t||_2^2 - ||a_{t+1} - z_t||_2^2 \right)$   
=  $\frac{1}{2} \left( ||a_t - z_t||_2^2 - ||(1 - \frac{1}{t}) \cdot (a_t - z_t)||_2^2 \right)$ 

• And from this,

$$(a_t, z_t) - (a_{t+1}, z_t) = \frac{1}{2} \left( ||a_t - z_t||_2^2 - (1 - \frac{1}{t})^2 \cdot ||a_t - z_t||_2^2 \right)$$
  
=  $\frac{1}{2} \left( 1 - (1 - \frac{1}{t})^2 \right) \cdot ||a_t - z_t||_2^2$   
=  $\left( \frac{1}{t} - \frac{1}{2t^2} \right) \cdot ||a_t - z_t||_2^2$   
 $\leq \frac{1}{t} \cdot ||a_t - z_t||_2^2.$ 

۱.

× × 0 × × ×

**Reminder:** 
$$(a_t, z_t) - (a_{t+1}, z_t) \le \frac{1}{t} \cdot ||a_t - z_t||_2^2$$
. (2)

• Since by assumption  $L = \sup_{z \in Z} ||z||_2$  and  $a_t$  is the empirical average of  $z_1, \ldots, z_{t-1}$ , we have that  $||a_t||_2 \le L$ .

**Reminder:** 
$$(a_t, z_t) - (a_{t+1}, z_t) \le \frac{1}{t} \cdot ||a_t - z_t||_2^2$$
. (2)

- Since by assumption  $L = \sup_{z \in \mathbb{Z}} ||z||_2$  and  $a_t$  is the empirical average of  $z_1, \ldots, z_{t-1}$ , we have that  $||a_t||_2 \le L$ .
- 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)

**Reminder:** 
$$(a_t, z_t) - (a_{t+1}, z_t) \le \frac{1}{t} \cdot ||a_t - z_t||_2^2$$
. (2)

- Since by assumption  $L = \sup_{z \in \mathbb{Z}} ||z||_2$  and  $a_t$  is the empirical average of  $z_1, \ldots, z_{t-1}$ , we have that  $||a_t||_2 \le L$ .
- 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

$$\sum_{l=1}^{T} \left( (a_l, z_l) - (a_{l+1}, z_l) \right) \le \sum_{l=1}^{T} \left( \frac{1}{t} \cdot ||a_l - z_l||_2^2 \right) \le \sum_{l=1}^{T} \frac{1}{t} \cdot (2L)^2$$
$$= 4L^2 \cdot \sum_{l=1}^{T} \frac{1}{t}.$$



X

××

$$\begin{array}{c} \textbf{Reminder:} \quad \sum_{t=1}^{T} \left( (a_t, z_t) - (a_{t+1}, z_t) \right) \leq 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t} \\ \textbf{\bullet} \text{ Now, it holds that } \sum_{t=1}^{T} \frac{1}{t} \leq \log(T) + 1, \text{ so that we obtain} \\ \\ \sum_{t=1}^{T} \left( (a_t, z_t) - (a_{t+1}, z_t) \right) \leq 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t} \leq 4L^2 \cdot \left( \log(T) + 1 \right), \end{array}$$

which is what we wanted to prove.

00X