FTL FOR OQO PROBLEMS

Advanced Machine Learning

Follow the leader for OQO problems

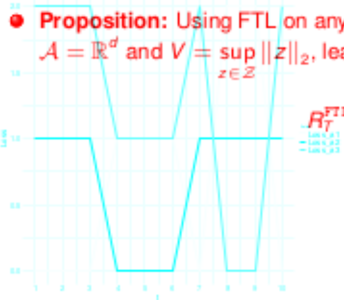
- One popular instantiation of the online learning problem is the problem of *online quadratic optimization* (OQO).
- In its most general form, the loss function is thereby defined as

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

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

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



Learning goals

- Prove that FTL works for online quadratic optimization problems

FTL FOR OQO PROBLEMS: ANALYSIS

- 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

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

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

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

- **Proof:**
  - In the following, we denote $a_1^{\mathrm{FTL}}, a_2^{\mathrm{FTL}}, \ldots$ simply by $a_1, a_2, \ldots$

FTL FOR OQO PROBLEMS: ANALYSIS

- **Proof:**
  - In the following, we denote $a_1^{\mathrm{FTL}}, a_2^{\mathrm{FTL}}, \ldots$ simply by $a_1, a_2, \ldots$

Reminder (Useful Lemma):

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

Reminder (Useful Lemma):

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

  - Using this lemma, we just have to show that

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

    Reminder (Useful Lemma):

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

  - Using this lemma, we just have to show that

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

  - 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.

FTL FOR OQO PROBLEMS: ANALYSIS

- Proof
- Claim: It holds that $a_t = \frac{1}{t-1} \cdot \sum_{s=1}^{t-1} z_s$, if $(a, z) = \frac{1}{2} \|a - z\|_2^2$.
  - In the following, we denote $a_1^{\mathrm{FTL}}, a_2^{\mathrm{FTL}}, \ldots$ simply by $a_1, a_2, \ldots$

Reminder (Useful Lemma):

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

  - Using this lemma, we just have to show that

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

  - 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.

FTL FOR OQO PROBLEMS: ANALYSIS

- 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

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

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

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

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

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

  - 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.)

FTL FOR OQO PROBLEMS: ANALYSIS

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

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

  - 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.)

- Hence, a_t is the empirical average of $z_1, \ldots, z_{t-1}$ and we can provide the following incremental update formula for its computation

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

FTL FOR OQO PROBLEMS: ANALYSIS

- Hence, a_t is the empirical average of $z_1, \ldots, z_{t-1}$ and we can provide the following incremental update formula for its computation

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

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

- From the last display we derive that

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

FTL FOR OQO PROBLEMS: ANALYSIS

- Hence, a_t is the empirical average of $z_1, \ldots, z_{t-1}$ and we can provide the following incremental update formula for its computation

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

- From the last display we derive that

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

- Claim:

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

- Indeed, this can be seen as follows

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

- Hence, a_t is the empirical average of $z_1, \dots, z_{t-1}$ and we can provide the following incremental update formula for its computation

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

- 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:

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

- Indeed, this can be seen as follows

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

- And from this,

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

- Indeed, this can be seen as follows

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

- And from this,

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

- Since by assumption $L = \sup_{z \in \mathcal{Z}} ||z||_2$ 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

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

FTL FOR OQO PROBLEMS: ANALYSIS

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

- Since by assumption $L = \sup_{z \in \mathcal{Z}} ||z||_2$ and a_t is the empirical average of $z_1, \ldots, z_{t-1}$, we have that $||a_t||_2 \leq L$.
- Now the triangle inequality states that for any two vectors $x, y \in \mathbb{R}^d$ it holds that

$$||x + y||_2 \leq ||x||_2 + ||y||_2 ,$$

so that

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

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

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

$$= 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t}.$$

FTL FOR OQO PROBLEMS: ANALYSIS

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

Reminder: $\sum_{t=1}^{T} (L(a_t, z_t) - L(a_{t+1}, z_t)) \leq 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t}$

- 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 $||a_t||_2 \leq L$.
- Now the triangle inequality states that for any two vectors $x, y \in \mathbb{R}^d$ it holds that

$$||x + y||_2 \leq ||x||_2 + ||y||_2 \,,$$

so that

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

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

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

$$= 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t}.$$

FTL FOR OQO PROBLEMS: ANALYSIS

Reminder: $\sum_{t=1}^{T} ((a_t, z_t) - (a_{t+1}, z_t)) \leq 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t}$

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

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

which is what we wanted to prove. □

FTL FOR OQO PROBLEMS: ANALYSIS

Reminder: $\sum_{t=1}^{T} (L(a_t, z_t) - L(a_{t+1}, z_t)) \leq 4L^2 \cdot \sum_{t=1}^{T} \frac{1}{t}$

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

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

which is what we wanted to prove. □