

Week 4 - Lecture 19

Lemma: Let w_1, w_2, \dots be the sequence produced by FTL.

Then $\forall u \in S$, we have

$$\begin{aligned} R(n, \text{FTL}, u) &= \sum_{t=1}^n f_t(w_t) - \sum_{t=1}^n f_t(u) \\ &\leq \sum_{t=1}^n f_t(w_t) - \sum_{t=1}^n f_t(w_{t+1}) \quad \leftarrow \text{To prove this} \quad \text{--- (1)} \end{aligned}$$

Showing this is same as showing

$$\sum_{t=1}^n f_t(w_{t+1}) \leq \sum_{t=1}^n f_t(u) \quad \text{--- (2)}$$

Proving by induction:

For $n=1$ ~~the~~ (2) holds by choice of w_2

Assume the inequality holds for $(n-1)$ rounds

$$\sum_{t=1}^{n-1} f_t(w_{t+1}) \leq \sum_{t=1}^{n-1} f_t(u)$$

add $f_n(w_{n+1})$ on both sides

$$\sum_{t=1}^n f_t(w_{t+1}) \leq f_n(w_{n+1}) + \sum_{t=1}^{n-1} f_t(u)$$

choose $u = w_{n+1}$ then:

$$\leq f_n(w_{n+1}) + \sum_{t=1}^{n-1} f_t(w_{n+1})$$

$$= \sum_{t=1}^n f_t(w_{t+1})$$

$$= \min_{u \in S} \sum_{t=1}^n f_t(u) \quad \leftarrow \text{from the algo}$$

$$\sum_{t=1}^n f_t(w_{t+1}) \leq \sum_{t=1}^n f_t(u) \quad \text{for any } u$$

Exploring the algorithm & bound for a convex fun

Taking a quadratic loss function

$$f_t = \frac{1}{2} \|w_t - v_t\|_2^2$$

In each round

$$w_t = \arg \min_w \sum_{i=1}^{t-1} \frac{1}{2} \|w - v_i\|_2^2$$

$$w_t = \frac{1}{t-1} \sum_{i=1}^{t-1} v_i$$

$$\begin{aligned} w_{t+1} &= \frac{1}{t} \sum_{i=1}^t v_i = \frac{(t-1)}{(t-1)t} \left(\sum_{i=1}^{t-1} v_i + v_t \right) \\ &= \left(1 - \frac{1}{t} \right) w_t + \frac{1}{t} v_t \end{aligned}$$

$$\Rightarrow w_{t+1} - v_t = \left(1 - \frac{1}{t} \right) (w_t - v_t) \quad \text{--- (3)}$$

Now plugging it back in eq (1)

$$f_t(w_t) - f_t(w_{t+1})$$

$$= \frac{1}{2} \|w_t - v_t\|_2^2 - \frac{1}{2} \|w_{t+1} - v_t\|_2^2$$

$$= \left(1 - \left(1 - \frac{1}{t} \right)^2 \right) \|w_t - v_t\|_2^2 \quad [\text{Plugging in values for (3)}]$$

$$= \frac{1}{2} \left(1 - \frac{1}{t} + \frac{1}{t} \right) \|w_t - v_t\|_2^2$$

$$= \frac{1}{2} \left(-\frac{1}{t^2} + \frac{2}{t} \right) \|w_t - v_t\|_2^2$$

$$\leq \frac{1}{t} \|w_t - v_t\|_2^2 \quad [\text{Ignoring } -1/t^2]$$

$$\text{Assume that } \|v_t\|_2^2 \leq L$$

Since w_t are averaging v_t

$$\therefore \|w_t\|_2^2 \leq L \quad [\text{not for } t=1, \text{ ignore that}]$$

$$\begin{aligned} \therefore \int_t(w_t) - \int_t(w_{t+1}) &\leq \frac{1}{t} \|w_t - v_t\|_2^2 \\ &= \frac{1}{t} (\|w_t\|_2^2 + \|v_t\|_2^2) \quad \left[\begin{array}{l} \text{Applying} \\ \text{triangular} \\ \text{inequality} \end{array} \right] \\ &= \frac{1}{t} (2L) \end{aligned}$$

Taking summation on both sides

$$\begin{aligned} \sum_{t=1}^n \int_t(w_t) - \sum_{t=1}^n \int_t(w_{t+1}) &\leq \sum_{t=1}^n \frac{1}{t} (2L) \\ &\leq 2L \log n \quad \text{--- (2)} \end{aligned}$$

Thus we have a bound that is independent of the u

$$\left[\int \frac{1}{x} dx = \log x \right]$$

Hence:

$$\underline{\underline{R(n, FTL) \leq 2L \log n}}$$

This bound holds for quadratic function, but does not hold for linear functions.

Hence a regularizer term needs to be added.

Next Lecture ↗