# Week 4 – Lecture 20

For convex functions of the form:

$$f_t(\omega) = \|\omega - z_t\|_2^2$$

$$R(n, FTL) \leq (2L)(\log n + 1)$$

$$\omega_t = \frac{1}{t-1} \sum_{i=1}^{t-1} z_i \quad \text{— (1)}$$

← FTL Algo for quadratic case

---

Let's take linear loss function

$f_t(\omega) = \langle \omega, z_t \rangle$ and check the bounds

Example: $S = [-1, 1]$

Taking only scalar values, $d = 1$

$$\therefore \; f_t(\omega) = \langle \omega, z_t \rangle = \omega z_t$$

Let
$$z_t = \begin{cases} 0.5 & \text{for } t = 1 \\ 1 & \text{if } t \text{ is even} \\ -1 & \text{if } t \text{ is odd and } t > 1 \end{cases}$$

Lets compute $\omega$ with FTL

$$\omega_t = \underset{\omega}{\operatorname{argmin}} \sum_{i=1}^{t} f_i(\omega)$$

$\omega_1 =$ ____ (anything can be played)

$$\omega_2 = \operatorname{argmin} (-0.5\,\omega)$$

$w \in S$

$$w_3 = \underset{w \in S}{\operatorname{argmin}} (-0.5w + w) = \underset{w \in S}{\operatorname{argmin}} (+0.5w) = -1$$

$$w_4 = \underset{w \in S}{\operatorname{argmin}} (-0.5w + w - w) = \underset{w \in S}{\operatorname{argmin}} (-0.5w) = 1$$

$$w_5 = -1$$

We see that the $w_i$ are alternating from one round to another.

Regret incurred

$$\sum_{t=1}^{n} f_t(w_t) - \sum f_t(u)$$

$$= f_1(w_1) + (n-1) - \sum_{t=1}^{n} z_t u$$

$$= 1 + (n-1) - \sum_{t=1}^{n} z_t u \qquad [\text{Let } f_1(w_1) = 1]$$

$$= n - u \sum_{t=1}^{n} z_t$$

~~to~~ minimizing over $u$

$$= n - \min_u u \sum_{t=1}^{n} z_t$$

$$= n + 0.5$$

or

$$= n - 0.5$$

If $n \to$ odd

$$\sum_{t=1}^{n} z_t = -0.5$$

If $n \to$ even

$$\sum_{t=1}^{n} z_t = 0.5$$

Thus regret is $\underline{\underline{O(n)}}$

In the quadratic case the $w_t$ was getting averaged hence not changing abruptly, thus updates were stable

In the linear case the $w_t$ was changing abruptly, henc[e] the updates were not stable; the past had no effect on the updates with ... future rounds.

Hence the solution would be to use a regularizer.

Follow the regularized leader Algorithm (FoReL)

$$\forall t, \; w_t = \underset{w \in S}{\operatorname{argmin}} \sum_{i=1}^{t-1} f_i(w) + \underbrace{R(w)}_{\text{regularizer}}$$

$\Rightarrow$ L2 Regularizer $\quad R(w) = \frac{1}{2\eta} \|w\|_2^2$

Let $f_t(w) = \langle w, z_t \rangle$

$$F_t = \sum_{i=1}^{t-1} f_t(w) + R(w)$$

$$= \sum_{i=1}^{t-1} \langle w, z_i \rangle + \frac{1}{2\eta} \|w\|_2^2$$

Differentiating $\underbrace{\text{and setting}}_{\text{w.r.t } w}$ $F'_t = 0$ (since minimizing)

$$F'_t = \sum_{i=1}^{t-1} z_i + \frac{1}{\eta} w$$

$$\therefore \; w^*_t = -\eta \sum_{i=1}^{t-1} z_i \quad \text{— } (2)$$

$$w^*_t = w^*_{t-1} - \eta z_{t-1} \quad \text{— } (2a)$$ [Itiratively writing (

This is similar to gradient descent

$z_{t-1}$ is the gradient of $f_t(w) = \langle w, z_t \rangle$

Hence the update rule is

$$w^*_t = w^*_{t-1} - \eta \nabla f_{t-1}(w) \longrightarrow \text{Online gradient descent}$$