Week 5 - Lecture 24

Lemma :- [When using FoReL]

$$f_t(\omega_t) - f_t(\omega_{t+1}) \leq \frac{L_t^2}{\sigma}$$

$\leftarrow$ Lipschitz constant for function f_t

$\uparrow$ σ- strongly convex for regularizer

Proof :-

$$\forall t, \quad F_t(\omega) = \sum_{i=1}^{t-1} f_i(\omega) + R(\omega)$$

$\downarrow$

this will be σ- strongly convex as:

$f_i(\omega)$ is convex

$R(\omega)$ is σ- strongly convex

When a σ- strongly convex function is added to a convex function then the resulting function becomes σ- strongly convex.

Update rule of FoReL:

$$\omega_t = \underset{\omega}{\operatorname{argmin}}\; F_t(\omega)$$

Since $F_t(\omega)$ is σ- strongly convex

$$\therefore \quad F_t(\omega_{t+1}) \geq F_t(\omega_t) + \frac{\sigma}{2} \|\omega_t - \omega_{t+1}\|^2 \quad \text{--- (i)}$$

$\nwarrow$

already discussed property of σ-strongly convex function

$$\therefore \; F_{t+1}(\omega_t) \geq F_{t+1}(\omega_{t+1}) + \frac{\sigma}{2} \|\omega_t - \omega_{t+1}\|^2 \quad \text{— (ii)}$$

Adding (i) & (ii)

$$F_{t+1}(\omega_t) - F_t(\omega_t) + F_t(\omega_{t+1}) - F_{t+1}(\omega_{t+1})$$

$$\geq \sigma \|\omega_t - \omega_{t+1}\|^2$$

$$\Rightarrow f_t(\omega_t) - f_t(\omega_{t+1}) \geq \sigma \|\omega_t - \omega_{t+1}\|^2 \quad \text{— (iii)}$$

[Using the defn of $F_t(\omega)$]

$$f_t(\omega_t) - f_t(\omega_{t+1}) \leq L_t \|\omega_t - \omega_{t+1}\| \quad \text{— (iv)}$$

[Applying the Lipschitzness of the function]

Thus we have lower & upper bound of

$$f_t(\omega_t) - f_t(\omega_{t+1})$$

Thus we have :

$$\sigma \|\omega_t - \omega_{t+1}\|^2 \leq L_t \|\omega_t - \omega_{t+1}\|$$

$$\Rightarrow \|\omega_t - \omega_{t+1}\| \leq L_t / \sigma \quad \text{— (v)}$$

Substituting this back in (iv), we have

$$f_t(w_t) - f_t(w_{t+1}) \leq L_t^2/\sigma$$

Thus Lemma proved.

Using this we can now write the Theorem:

Theorem: - Let $f_1\, f_2 \ldots f_n$ are convex function s.t. f_t is L_t-Lipschitz w.r.t. $\|\cdot\|$.

Then if FoReL is used with a σ-strongly convex w.r.t. same norm then:

$$\text{Regret}(n,u) \leq R(u) - \min_{v \in S} R(v) + \frac{nL^2}{\sigma}$$

This is because:

We know:

$$\text{Regret}(n,u) \leq R(u) - R(w_1) + \sum_{t=1}^{n} f(w_t) - f(w_{t+1})$$

This is equivalent to $\min_{w} R(w)$

from the Lemma we have:

$$\sum_{t=1}^{n} f(w_t) - f(w_{t+1}) = \sum_{t=1}^{n} L_t^2/\sigma$$

Now if f_ts have the same Lipschitz constant then :

$$= nL^2/\sigma$$

However if they are different

$$= \sum_{t=1}^{n} L_t^2 / \sigma$$

$$= n\left(\frac{\sum L_t^2}{n}\right) / \sigma$$

Multiplying & dividing by n

If we assume the $\frac{\sum L_t^2}{n} = L^2$, then :

$$= nL^2/\sigma$$

this needs to be included in the theorem when stating it

Regularizers :

1) $R(u) = \frac{1}{2}\|w\|_2^2$ for $\forall u \in \mathbb{R}^d$

This is 1-strongly convex with l_2-norm

To check this :

$$\langle \nabla^2 R(\omega) x, x \rangle \geqslant \sigma \|x\|^2$$

Now $R(u) = \frac{1}{2\eta}\|w\|_2^2$

is $\frac{1}{\eta}$-strongly convex with l_2-norm

2) $R(u) = \sum_{i=1}^{d} u_i \log u_i$

$u \in S = \{w,\ w_i > 0,\ \underset{\substack{\uparrow \\ l_1\text{-norm}}}{\sum w_i = 1}\}$ (probability space)

$\therefore$ Entropy regularizer is 1-strongly convex w.r.t l_1-norm

if however $u \in S = \{w, w_i > 0, \sum w_i \leq B\}$

then it is $\frac{1}{B}$-strongly convex w.r.t l_1-norm

Now <u>to find</u> : the bounds when using these regularizers.