Assignment 5

Q.1 : Given : Linear convex optimization problem where $f_t(\omega) = \langle \omega_t, z_t \rangle$ where $z_t \in [0,1]^d \; \forall t$ and $S = \{\omega : \|\omega\|_1 = 1, \omega > 0\}$.

Update rule for Weighted Majority :

$$\omega_i^{t+1} = \omega_i^t \exp(-\eta z_{ti}) \quad \text{for the i-th expert} \qquad —— (1)$$

FoReL update Rule :

$$\forall t, \; \omega_t = \underset{\omega \in S}{\operatorname{argmin}} \sum_{i=1}^{t-1} f_i(\omega) + R(\omega) \qquad —— (2)$$

Entropy Regularizer :

$$R(\omega) = \frac{1}{\eta} \sum_{k=1}^{d} \omega_k \log \omega_k$$

Thus FoReL update rule with Entropy Regularizer

$$\omega_t = \underset{\omega \in S}{\operatorname{argmin}} \sum_{i=1}^{t-1} f_i(\omega) + \frac{1}{\eta} \sum_{k=1}^{d} \omega_k \log \omega_k \qquad —— (4)$$

Given $f_t(\omega) = \langle \omega_t, z_t \rangle$

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

$$= \sum_{i=1}^{t-1} \langle w, z_i \rangle + \frac{1}{\eta} \sum_{k=1}^{d} w_k \log w_k$$

Differentiating w.r.t. w

$$F_t' = \sum_{i=1}^{t} z_i + \frac{1}{\eta} \sum_{k=1}^{d} [\log(w_k) + 1]$$

$$F_t' = \sum_{i=1}^{t} z_i + \frac{d}{\eta} \sum_{k=1}^{d} \log(w_k)$$

Setting $F_t' = 0$ [as $w = \operatorname{argmin} F_t$]

$$\Rightarrow 0 = \sum_{i=1}^{t} z_i + \frac{d}{\eta} \sum_{k=1}^{d} \log w_k$$

$$\Rightarrow \sum_{k=1}^{d} \log(w_k) = -\frac{\eta}{d} \sum_{i=1}^{t} z_i$$

Considering for one arm at round t, we have:

$$\log(w_t) = -\eta z_{tk}$$

$$\Rightarrow w_t = \exp(-\eta z_{tk})$$

$$\Rightarrow \underline{w_t = w_{t-1} \exp(-\eta z_{(t-1)k})} \quad \text{——} \quad (5)$$

Hence the update rule for FoReL with

Entropy Regularizer is equivalent to the update rule of weighted majority.

Regret of Weighted Majority

$$\text{Regret}(n) = \sum_{t=1}^{n} \langle w_t, z_t \rangle - \min_{i \in [d]} \sum_{t=1}^{n} V_{ti}$$

$$\leq \sqrt{2n \log d} \qquad \text{——— (6)}$$

We know the Theorem :

Let $f_1, \dots f_n$ be a sequence of convex functions s.t. f_t is L_t-Lipschitz with respect to norm $\|\cdot\|$.
Let L be such that $\frac{1}{n} \sum_{t=1}^{n} L_t^2 \leq L^2$.

If FoReL is run on the sequence with a regularizer which is σ-strongly convex w.r.t. the same norm. Then, $\forall\, u \in S$,

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

$$\text{——— (7)}$$

Thus for FoReL with Entropy Regularizer

[Entropy Regularizer is $\frac{1}{\eta}$ - strongly convex]

$$\text{Regret}(n,u) \leq R(u) - \min_{u \in S} R(u) + n\eta L^2$$

$$\leq R(u) + \max(-R(u)) + n\eta L^2$$

$$\boxed{\begin{aligned} -R(u) &= \frac{1}{\eta}\sum_{i=1}^{d} u_i \log \frac{1}{u_i} \\ &\leq \frac{1}{\eta} \log d \sum_{i=1}^{d} 1/u_i \\ &\leq \frac{1}{\eta} \log d \end{aligned}}$$

$$\leq R(u) + \frac{1}{\eta}\log d + n\eta L^2$$

Now Given that let U be a set of unit vectors

Thus $U \in \{e_1, e_2, \dots e_d\}$ where

$$e_1 = \{1, 0, 0, \dots 0\}$$
$$e_2 = \{0, 1, 0, \dots 0\}$$
$$\vdots$$
$$e_d = \{0, 0, 0, \dots d\}$$

Now considering u to be any e_i, we have:

$$\frac{1}{\eta} \sum_{i=1}^{d} u_i \log u_i = 0$$

Thus the Regret simplifies to:

$$\text{Regret}(n, u) \leq \frac{1}{\eta} \log d + L^2 n \eta$$

Setting $\eta = \frac{\sqrt{\log d}}{L\sqrt{2n}}$ we get

$$\text{Regret}(n, u) \leq L\sqrt{2n \log d} \quad \text{—— } (8)$$

For prediction with expert advice $L = 1$

$$\therefore \text{Regret}(n, \text{FoReL}, u) \leq \sqrt{2n \log d} \quad \text{—— } (9)$$

which is similar to the regret bound obtained for weighted majority.

Q.2: Given: For any differentiable function $R : \mathbb{R}^d \to \mathbb{R}$, for vectors w and u

Bregman Distance is defined as:

$$D_R(w \| u) = R(w) - R(u) - \langle \nabla R(u), w - u \rangle$$

$- \quad " \quad "^2$

To find: The Bregman divergance of $R = \frac{1}{2}\|u\|_2$

$$R = \frac{1}{2}\|u\|_2^2$$

$$\nabla R(u) = u$$

$$D_R(w\|u) = \frac{1}{2}\|w\|_2^2 - \frac{1}{2}\|u\|_2^2 - \langle u, w-u \rangle$$

$$= \frac{1}{2}\sum_i w_i^2 + \frac{1}{2}\sum_i u_i^2 - \sum_i \left(\langle w,u \rangle - \langle u,u \rangle\right)$$

$$= \frac{1}{2}\sum_i \langle w,w \rangle + \frac{1}{2}\sum_i \langle u,u \rangle - \sum_i \langle w,u \rangle + \sum_i \langle u,u \rangle$$

$$= \frac{1}{2}\sum_i \langle w,w \rangle - \sum \langle w,u \rangle + \frac{1}{2}\sum_i \langle u,u \rangle$$

$$\Rightarrow D_R(w\|u) = \frac{1}{2}\|w-u\|_2^2$$