Assignment 4
Question 1:

Assignment 4

Q.1.

Prove for any ω, z

$$\langle \omega, z \rangle \leq \|\omega\| \|z\|_*$$

If $\|x\|$ is a norm, it's dual norm $\|z\|_*$ is

$$\|z\|_* = \sup u^T z \qquad — ①$$

where u represents a vector such that $\|u\| \leq 1$

For any two vectors, ω and z, inner product can be represented as

$$\langle \omega, z \rangle = \omega^T * z$$ where $*$ represents element-wise product.

$$\omega^T z = \|\omega\| \frac{\omega^T z}{\|\omega\|} \qquad — ②$$

$\frac{\omega^T}{\|\omega\|}$ is a unit vector or we can write that :- $\left\| \frac{\omega^T}{\|\omega\|} \right\| = 1$

Therefore we can say that u satisfies this condition and hence let

$$\|u^T\| = \left\| \frac{\omega^T}{\|\omega\|} \right\| \leq 1 \qquad — ③$$

Substituting ③ in ②, we get

$$\omega^T z = \|\omega\| \; u^T z \qquad — ④$$

Since u^T is a unit vector, therefore from ①

$$u^T z \leq \sup u^T z = \|z\|_* \qquad — ⑤.$$

Putting in eq. ④,

$$\omega^T z = \|\omega\| \; u^T z \leq \|\omega\| \|z\|_*$$

$$\omega^T z \leq \|\omega\| \|z\|_*$$

$$\langle \omega, z \rangle \leq \|\omega\| \|z\|_*$$

Hence proved.

Question 2 :

The regret of A on each period of 2^m rounds is upper bounded by :-

$$R_m \leq \alpha\sqrt{2^m}$$

Therefore, the total regret if we stop after m rounds is given by summing R_m for each round.

$$\sum_{m=1}^{\lceil \log_2(m) \rceil} R_m = \text{Total Regret after } m \text{ rounds}$$

$$= \sum_{m=1}^{\lceil \log_2(m) \rceil} \alpha\sqrt{2^m}$$

$$= \alpha \sum \sqrt{2} + \sqrt{2^1} + \sqrt{2^2} + \dots + \sqrt{2^{\lceil \log_2(m) \rceil}}$$

$$= \alpha \frac{1 - \sqrt{2}^{\log_2(m)+1}}{1 - \sqrt{2}}$$

$$= \alpha \frac{1 - \sqrt{2}\left(2^{\frac{1}{2}\log_2(m)}\right)}{1 - \sqrt{2}}$$

$$= \alpha \frac{1 - \sqrt{2m}}{1 - \sqrt{2}}$$

$$R = \alpha \frac{\sqrt{2}}{\sqrt{2}-1} \frac{\sqrt{2m} - 1}{\sqrt{2}}$$

$$R \leq \frac{\sqrt{2}}{\sqrt{2}-1} \alpha \sqrt{m}$$

Hence proved.

Question 3:

Q.3 We know that for OGD:

Input $\eta > 0$

initialise $\theta_1 = 0$

for $t = 0, 1, 2, \ldots$

update rule : $\theta_{t+1} = \theta_t - \eta z_t$

$= \theta_t - \eta \partial f_t(\theta_t)$

since $z_t \in \partial f_t(\theta_t)$

Given for online Mirror Descent :-

Input : A link function defined as

$g : R^d \longrightarrow S$

such that $g(\theta) = \eta\theta$

Initialize : $\theta_1 = 0$

for $t = 0, 1, 2, \ldots$

predict $w_t = \eta\theta_t$

update rule $\theta_{t+1} = \theta_t - z_t$ where $z_t \in \partial f_t(w_t)$

$= \theta_t - \partial f_t(w_t)$

$= \theta_t - \partial f_t(\eta\theta_t)$

$= \theta_t - \eta \partial f_t(\theta_t)$

Hence, by taking $g(\theta) = \eta\theta$, the same update rule is obtained as the online gradient descent.