ML HW

Q2, Linear Regression model $\vec{y} = \vec{X}\vec{\theta} + \epsilon$ where $\vec{y} \in \mathbb{R}^n$, $\vec{X} \in \mathbb{R}^{n\times d}$, $\vec{\theta} \in \mathbb{R}^d$, $\epsilon \in \mathbb{R}^n$

(a) Find general form of θ^* that minimize the weighted MSE

Given: $L(\theta) = (y - X\theta)^T \Omega (y - X\theta)$

$$= (y^T - \theta^T X^T)\Omega(y - X\theta) = (y^T\Omega - \theta^T X^T \Omega)(y - X\theta)$$

$$= y^T\Omega y - \underbrace{\theta^T X^T \Omega y}_{③} - \underbrace{y^T \Omega X\theta}_{②} + \underbrace{\theta^T X^T \Omega X\theta}_{①} \rightarrow \text{2nd order term}$$

$$= (\theta - \phi)^T X^T \Omega X (\theta - \phi) + \text{complementary term}$$

shorted as CT

$$= (\theta^T - \phi^T)(X^T\Omega X)(\theta - \phi) + CT$$

$$= (\theta^T X^T \Omega X - \phi^T X^T \Omega X)(\theta - \phi) + CT$$

$$= \underbrace{\theta^T X^T \Omega X\theta}_{\triangle 1} - \underbrace{\phi^T X^T \Omega X\theta}_{\triangle 2} - \underbrace{\theta^T X^T \Omega X\phi}_{\triangle 3} + \phi^T X^T \Omega X\phi + CT$$

the same as the last term

$$\left.\begin{aligned} \triangle 1 &= ① \\ \triangle 2 &= ② \\ \triangle 3 &= ③ \end{aligned}\right\} \quad \theta^T X^T \Omega y = \theta^T X^T \Omega X\phi \Rightarrow \phi = (X^T\Omega X)^{-1} X^T \Omega y$$

put into $\triangle 2 = ② \Rightarrow \theta^T X^T \Omega X\theta = \theta^T \cancel{X^T\Omega X (X^T \Omega X)^{-1}} X^T \Omega y = \theta^T X^T \Omega y$

the same with ②

$$CT = y^T\Omega y - \phi^T X^T \Omega X \phi = y^T \Omega y - y^T \Omega^T X\left[(X^T\Omega X)^{-1}\right]^T \cancel{X^T\Omega X (X^T\Omega X)^{-1}} X^T \Omega y$$

$$= y^T\Omega y - y^T\Omega^T X\left[(X^T\Omega X)^{-1}\right]^T X^T \Omega y$$

$$\boxed{\begin{aligned} L(\theta) &= (\theta - \phi)^T X^T \Omega X (\theta - \phi) + y^T\Omega y - y^T \Omega^T X\left[(X^T\Omega X)^{-1}\right]^T X^T\Omega y \\ &\text{where } \phi = (X^T\Omega X)^{-1} X^T \Omega y \end{aligned}}$$

\#