2017831003

Mahfuzur Rahman Emon

Machine learning Assignment - 1

Question : 1

Here, $f(z) = \log_e (1+z)$

where, $z = x^T x$, $x \in R^d$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$x^T = \begin{bmatrix} x_1 & x_2 & \dots & x_d \end{bmatrix}$$

$$\therefore x^T x = \left[x_1^2 + x_2^2 + \dots + x_d^2 \right]$$

Applying chain rule,

$$\frac{df}{dx} = \frac{df}{dz} \cdot \frac{dz}{dx}$$

$$= \frac{d}{dz} \left(\log_e (1+z) \right) \cdot \frac{d}{dx} \left(x^T x \right)$$

$$= \frac{d}{dz} \left(\log_e (1+z) \right) \cdot \frac{d}{dx} \left(x_1^2 + x_2^2 + \dots + x_d^2 \right)$$

$$= \frac{1}{1+z} \frac{d}{dz}(z) \cdot (2x_1 + 2x_2 + \cdots + 2x_d)$$

$$= \frac{1}{1+z} \cdot 2 \, (x_1 + x_2 + \cdots + x_d)$$

$$= \frac{2}{1+z} \sum_{i=1}^{d} x_i$$

$\therefore$ The gradient ~~of~~ f is $\frac{2}{1+z} \cdot \sum_{i=1}^{d} x_i$

Question 2 :

Here, $f(z) = e^{-\frac{z}{2}}$, $z = g(y) = y^T S^{-1} y$

$y = h(x) = x - \mu$

where, $x, \mu \in R^d$ and $S \in R^{d \times d}$

Using the chain Rule,

$$\frac{df}{dx} = \frac{df}{dz} \times \frac{dz}{dy} \times \frac{dy}{dx}$$

$$\therefore \frac{df}{dz} = \frac{d}{dz}\left(e^{-z/2}\right) = -\frac{e^{-z/2}}{2}$$

$$\frac{dz}{dy} = \frac{d}{dy}\left(y^T S^{-1} y\right)$$

$$= \lim_{h \to 0} \frac{g(y+h) - g(y)}{h}$$

$$= \lim_{h \to 0} \frac{(y^T + h)\, S^{-1} (y+h) - y^T S^{-1} y}{h}$$

$$= \lim_{h \to 0} \frac{(y^T S^{-1} + h S^{-1})(y+h) - y^T S^{-1} y}{h}$$

$$= \lim_{h \to 0} \frac{\cancel{y^T S^{-1} y} + y^T S^{-1} h + h S^{-1} y + h^2 S^{-1} - \cancel{y^T S^{-1} y}}{h}$$

$$= \lim_{h \to 0} \frac{y^T S^{-1} h + h S^{-1} y + h^2 S^{-1}}{h}$$

$$= \lim_{h \to 0} \frac{h\,(y^T S^{-1} + S^{-1} y + h S^{-1})}{h}$$

$$= \lim_{h \to 0} \left(y^T S^{-1} + S^{-1} y + S^{-1} h\right)$$

$$= y^T S^{-1} + S^{-1} y + \lim_{h \to 0} (S^{-1} h)$$

$$= y^T S^{-1} + S^{-1} y$$

$$\therefore \frac{dy}{dx} = \frac{d}{dx}(x - \mu) = 1$$

$$\therefore \frac{df}{dx} = \frac{df}{dz} \times \frac{dz}{dy} \times \frac{dy}{dx}$$

$$= -\frac{e^{-z/2}}{2} \cdot \left(y^T S^{-1} + S^{-1} y\right) \cdot 1$$

$$= -\frac{e^{-z/2}}{2} \cdot \frac{1}{S}\left(y^T + y\right).$$

$\therefore$ The gradient is $-\dfrac{e^{-z/2}}{2} \cdot \dfrac{1}{S} \cdot \left(y^T + y\right)$