Registration No: 2019831040
Name: Jisan Ahmed

Mechine Learning Assignment No: 01

Qustion 1: $f(z) = \log_e(1+z)$

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \quad ; \quad x^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_d \end{bmatrix}$$

$$\therefore x^T x = [x_1^2 + x_2^2 + x_3^2 + \ldots + x_d^2]$$

Applying chain rule with respect to x

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

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

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

$$= \frac{1}{1+z} \cdot (2x_1 + 2x_2 + 2x_3 + \dots + 2x_d)$$

$$= \frac{2}{1+z} (x_1 + x_2 + x_3 + \dots + x_d)$$

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

Therefore, the gradient is $\frac{2}{1+z} \sum_{i=1}^{d} x_i$

Question 2:

Here $f(z) = e^{-\frac{z}{2}}$,

$$z = g(y) = y^T S^{-1} y$$

$$y = h(x) \cancel{= x - \mu}$$

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

here, $\mu \in \mathbb{R}^d$, $S \in \mathbb{R}^{d^2}$

Using the chain rule:

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

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

$$= -\frac{e^{-\frac{z}{2}}}{2}$$

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

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

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

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

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

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

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

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

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

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

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

Therefore, the gradient is $-\frac{e^{-\frac{z}{2}}}{2} \cdot \frac{1}{s}\left(y^T + y\right)$

CamScanner