Gourab Saha
2017831004

ML Assignment-1

Compute the gradient (First derivative) of the following function.

a) $f(z) = \log_e(1+z)$

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

b) $f(z) = \exp(-\frac{1}{2} z)$

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

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

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

Answer to the Question (a):

(a) $f(z) = \log_e(1+z)$

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_d \end{bmatrix} \qquad x^T = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_d \end{bmatrix}$$

$$x^T x = [x_1 \; x_2 \; x_3 \; \cdots \; x_d] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ \cdots \\ x_d \end{bmatrix}$$

$$= [x_1^2 + x_2^2 + x_3^2 + \cdots + x_d^2]$$

Using chain rule,

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

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

$$= \left(\frac{1}{1+z}\right) \cdot \frac{d}{dx}[x_1^2 + x_2^2 + x_3^2 + \cdots + x_d^2]$$

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

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

(Ans)

Answer to the question (b):

(b) $f(z) = \exp(-\frac{1}{2} z)$

$= e^{-\frac{1}{2} z}$

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

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

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

Using chain Rules,

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

$$= \frac{d}{dz}(e^{-\frac{z}{2}}) \cdot \frac{d}{dy}(y^T S^{-1} y) \cdot \frac{d}{dx}(x - \mu)$$

Computing the derivatives of

$\frac{d}{dz}(e^{-\frac{z}{2}})$, $\frac{d}{dy}(y^T S^{-1} y)$, $\frac{d}{dx}(x - \mu)$

(i) $\frac{d}{dz}(e^{-\frac{z}{2}}) = -\frac{1}{2} e^{-\frac{z}{2}}$

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

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

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

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

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

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

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

So,

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

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

So, the gradient is, $-\frac{e^{-\frac{z}{2}}}{2} \cdot (y^T S^{-1} + S^{-1} y)$

(Ans)