EECS 495 HW1

Liu Cao

Stu ID: 3072379

Excercises 2.1 Practice derivative calculations I

a) $g(w) = \frac{1}{2}qw^2 + rw + d$

First derivative: $g'(w) = \frac{1}{2}q \cdot 2w + r = qw + r$

Second derivative: $g''(w) = q$

b) $g(w) = -\cos(2\pi w^2) + w^2$

First derivative: $g'(w) = \sin(2\pi w^2) \cdot 2\pi \cdot 2w + 2w$

$= 4\pi w \sin(2\pi w^2) + 2w$

Second derivative: $g''(w) = 4\pi \sin(2\pi w^2) + 4\pi w \cos(2\pi w^2) \cdot 2\pi \cdot 2w + 2$

$= 4\pi \sin(2\pi w^2) + 16\pi^2 w^2 \cos(2\pi w^2) + 2$

c) $g(w) = \sum_{p=1}^{P} \log(1 + e^{-a_p w})$

First derivative: $g'(w) = \sum_{p=1}^{P} \frac{-a_p e^{-a_p w}}{1 + e^{-a_p w}} = \sum_{p=1}^{P} \frac{-a_p}{e^{a_p w} + 1}$

Second derivative: $g''(w) = \sum_{p=1}^{P} \frac{a_p \cdot e^{a_p w} \cdot a_p}{(e^{a_p w} + 1)^2} = \sum_{p=1}^{P} \frac{a_p^2 e^{a_p w}}{(e^{a_p w} + 1)^2}$

Excercises 2.2 Practice derivative calculations II.

a) $g(\bar{w}) = \frac{1}{2}\bar{w}^T \bar{Q} \bar{w} + \bar{r}^T \bar{w} + d$

$= \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} w_n Q_{nm} w_m + \sum_{n=1}^{N} r_n w_n + d$

$\therefore \frac{\partial g(\bar{w})}{\partial w_j} = \frac{1}{2}\left(\sum_{n=1}^{N} w_n Q_{nj} + \sum_{m=1}^{N} Q_{jm} w_m\right) + r_j$

$\because$ $\bar{Q}$ is $N \times N$ symmetric matrix

Gradient $\therefore$ $\nabla g(\bar{w}) = \frac{1}{2}(\bar{Q} + Q^T)\bar{w} + \bar{r}$

$= \bar{Q}\bar{w} + \bar{r}$

$\therefore \frac{\partial g(\bar{w})}{\partial w_j \partial w_i} = \frac{1}{2}(Q_{ij} + Q_{ji})$

Hessian: $\therefore$ $\nabla^2 g(\bar{w}) = \frac{1}{2}(\bar{Q} + \bar{Q}^T) = \bar{Q}$

b) $g(\bar{w}) = -\cos(2\pi \bar{w}^T \bar{w}) + \bar{w}^T \bar{w}$

$$= -\cos\left(2\pi \sum_{n=1}^{N} w_n^2\right) + \sum_{n=1}^{N} w_n^2$$

$$\therefore \frac{\partial g(\bar{w})}{\partial w_j} = \sin\left(2\pi \sum_{n=1}^{N} w_n^2\right) 4\pi w_j + 2w_j.$$

Gradient: $\therefore \nabla g(\bar{w}) = \sin(2\pi \bar{w}^T \bar{w}) 4\pi \bar{w} + 2\bar{w}$

$$\therefore \frac{\partial^2 g(\bar{w})}{\partial w_i \partial w_j} = \begin{cases} \cos\left(2\pi \sum_{n=1}^{N} w_n^2\right)(4\pi)^2 w_i w_j + \sin\left(2\pi \sum_{n=1}^{N} w_n^2\right) 4\pi + 2 & \text{if } i=j \\ \cos\left(2\pi \sum_{n=1}^{N} w_n^2\right)(4\pi)^2 w_i w_j & \text{if } i \neq j \end{cases}$$

Denote $\bar{I}_{N\times N}$ is the $N \times N$ identity matrix

$\therefore$ Hessian: $\nabla^2 g(\bar{w}) = \cos(2\pi \bar{w}^T \bar{w})(4\pi)^2 \bar{w}\bar{w}^T + (2 + \sin(2\pi \bar{w}^T \bar{w}) 4\pi) \bar{I}_{N\times N}$

c) $g(\bar{w}) = \sum_{p=1}^{P} \log(1 + e^{-\bar{a}_p^T \bar{w}})$

Denote $h_p(\bar{w}) = \log(1 + e^{-\bar{a}_p^T \bar{w}}) = \log\left(1 + e^{-\sum_{n=1}^{N} a_{pn} w_n}\right)$.

$$\therefore \frac{\partial}{\partial w_j} h_p(\bar{w}) = \frac{-e^{-\sum_{n=1}^{N} a_{pn} w_n} \cdot a_{pj}}{1 + e^{-\sum_{n=1}^{N} a_{pn} w_n}} = \frac{-e^{-\bar{a}_p^T \bar{w}} \cdot a_{pj}}{1 + e^{-\bar{a}_p^T \bar{w}}}$$

$$= \frac{-a_{pj}}{1 + e^{\bar{a}_p^T \bar{w}}}$$

$$\therefore \frac{\partial}{\partial w_j} g(\bar{w}) = -\sum_{p=1}^{P} \frac{a_{pj}}{1 + e^{\bar{a}_p^T \bar{w}}}$$

Gradient: $\therefore \nabla g(\bar{w}) = -\sum_{p=1}^{P} \frac{\bar{a}_p}{1 + e^{\bar{a}_p^T \bar{w}}}$

$$\therefore \frac{\partial^2 g(\bar{w})}{\partial w_i \partial w_j} = \sum_{p=1}^{P} \frac{a_{pi} \cdot e^{\bar{a}_p^T \bar{w}} \cdot a_{pj}}{(1 + e^{\bar{a}_p^T \bar{w}})^2}$$

$\therefore$ Hessian: $\nabla^2 g(\bar{w}) = \sum_{p=1}^{P} \frac{e^{\bar{a}_p^T \bar{w}}}{(1 + e^{\bar{a}_p^T \bar{w}})^2} \bar{a}_p \cdot \bar{a}_p^T$

Exercises 2.5 First order Taylor series geometry

We need to prove the vector on the hyperplane perpendicular to the normal vector.

According to Equation (2.3), we can know $(\bar{v}, g(\bar{v}))$ is on the hyperplane. we need find a another point on the hyperplane (which is close to $(\bar{v}, g(\bar{v}))$ so that we can use Taylor series approximation).

Suppose another point is $(\bar{w}, h(\bar{w}))$ The vector on the hyperplane $\vec{m} = (g(\bar{v}) - h(\bar{w}), \bar{v} - \bar{w})$ $\therefore$ we just prove $\vec{m} \cdot \vec{n} = 0$.

~~Suppose $\|\nabla g(\bar{v})\|$~~

$\therefore \vec{n} \cdot \vec{m} = (g(\bar{v}) - h(\bar{w})) \cdot 1 - \nabla g(\bar{v}) \cdot (\bar{v} - \bar{w})$

$\because \nabla g(\vec{v}) = \nabla g(\bar{v})^T \quad = (g(\bar{v}) - g(\bar{v}) - \nabla g(\bar{v})^T (\bar{w} - \bar{v})) - \nabla g(\bar{v})(\bar{v} - \bar{w})$

$= 0$

$\therefore$ The normal vector is : $\bar{n} = \begin{bmatrix} 1 \\ -\nabla g(\bar{v}) \end{bmatrix}$.

Exercise 2.7 Second order convexity calculations

a) $g(w) = w^2$, $g'(w) = 2w$

$g''(w) = 2 > 0$

$\therefore$ It is convex function.

b) $g(w) = e^{w^2}$

$g'(w) = e^{w^2} \cdot 2w$

$g''(w) = 4e^{w^2} w^2 + 2e^{w^2} > 0$.

$\therefore$ It is convex function.

(c) $g(w) = \log(1 + e^w)$

$$g'(w) = \frac{e^w}{1 + e^w}$$

$$g''(w) = \frac{e^w(1+e^w) - e^w \cdot e^w}{(1+e^w)^2} = \frac{e^w}{(1+e^w)^2} > 0$$

$\therefore$ It is convex function.

d) $g(w) = -\log(w)$

$g'(w) = -\frac{1}{w}$

$g''(w) = \frac{1}{w^2} > 0$

$\therefore$ It is a convex function