EE 475 HW1

Name: Zhisheng Liu
student #: 3226299

Exercise 2.1

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

$g'(w) = qw + r$

$g''(w) = q$

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

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

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

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

$g'(w) = \sum_{p=1}^{P} \frac{d}{dw}\left[\log(1 + e^{-a_p w})\right]$

$= \sum_{p=1}^{P} \frac{e^{-a_p w} \cdot (-a_p)}{1 + e^{-a_p w}}$

$g''(w) = \sum_{p=1}^{P} \frac{\left[e^{-a_p w} \cdot (-a_p)^2\right] \cdot (1 + e^{-a_p w}) - \left[e^{-a_p w} \cdot (-a_p)\right] \cdot \left[e^{-a_p w} \cdot (-a_p)\right]}{(1 + e^{-a_p w})^2}$

$= \sum_{p=1}^{P} \frac{a_p^2 e^{-a_p w} \cdot (1 + e^{-a_p w}) - e^{-2a_p w} \cdot a_p^2}{(1 + e^{-a_p w})^2}$

$= \sum_{p=1}^{P} \frac{a_p^2 \cdot e^{-a_p w}}{(1 + e^{-a_p w})^2}$

Exercise 2.2

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

$\Rightarrow g(w) = \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} w_n Q_{mn} w_n + \sum_{n=1}^{N} r_n w_n + d$

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

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

$\Rightarrow \nabla g(w) = QW + r$

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

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

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

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

$\Rightarrow \nabla g(w) = \sin(2\pi W^T W) \cdot 4\pi W + 2W$

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

$$\Rightarrow \frac{\partial^2 g(w)}{\partial w_i \partial w_j} = \begin{cases} \cos(2\pi W^T W)(4\pi)^2 \cdot WW^T + \sin(2\pi W^T W) \cdot 4\pi + 2 & , i = j \\ \cos(2\pi W^T W)(4\pi)^2 \cdot WW^T & , i \neq j \end{cases}$$

$\therefore \nabla^2 g(w) = \cos(2\pi w^T w) \cdot (4\pi)^2 WW^T + [\sin(2\pi W^T W) \cdot 4\pi + 2] \cdot I_{N \times N}$

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

$\because \frac{\partial g(w)}{\partial w_j} = \sum_{p=1}^{P} \frac{\partial}{\partial w_j} \log(1 + e^{-a_p^T w})$

$= \sum_{p=1}^{P} \frac{\partial}{\partial w_j} \log\left(1 + e^{-\sum_{n=1}^{N} a_{pn} w_n}\right)$

$= \sum_{p=1}^{P} \frac{e^{-\sum_{n=1}^{N} a_{pn} w_n} \cdot (-a_{pj})}{\left(1 + e^{-\sum_{n=1}^{N} a_{pn} w_n}\right)}$

$\therefore \nabla g(w) = \sum_{p=1}^{P} \frac{e^{-a_p^T W} \cdot (-a_p)}{1 + e^{-a_p^T W}}$

Name : Zhisheng Lin
student # : 3226299

$$\because \frac{\partial}{\partial w_i \partial w_j} g(w) = \sum_{p=1}^{P} \frac{\partial}{\partial w_i} \frac{e^{-\sum_{n=1}^{N} a_{pn} w_n} \cdot (-a_{pj})}{1+e^{-\sum_{n=1}^{N} a_{pn} w_n}}$$

$$= \sum_{p=1}^{P} \frac{e^{-\sum_{n=1}^{N} a_{pn} w_n} \cdot a_{pi} \cdot a_{pj}}{(1+e^{-\sum_{n=1}^{N} a_{pn} w_n})^2}$$

$$\therefore \nabla^2 g(w) = \sum_{p=1}^{P} \frac{e^{-a_p^T w} \cdot a_p \cdot a_p^T}{(1+e^{-a_p^T w})^2}$$

Exercise 2.5.

Since the vector in the tangent hyperplane can be represented as $[h(w)-h(v) \quad w-v]$

$$\because n^T = \begin{bmatrix} 1 & -\nabla g(v)^T \end{bmatrix}$$

$$\therefore n^T \cdot \begin{bmatrix} h(w)-h(v) \\ w-v \end{bmatrix} = \begin{bmatrix} 1 & -\nabla g(v)^T \end{bmatrix} \begin{bmatrix} h(w)-h(v) \\ w-v \end{bmatrix} = [h(w)-h(v)] - \nabla g(v)^T (w-v)$$

$$\because g(v) = h(v) \text{ at } w=v \quad \text{and } h(w) = g(v) + \nabla g(v)^T (w-v)$$

$$\therefore h(w)-h(v) = h(w)-g(v) = \nabla g(v)^T (w-v)$$

$$\therefore n^T \cdot \begin{bmatrix} h(w)-h(v) \\ w-v \end{bmatrix} = \nabla g(v)^T (w-v) - \nabla g(v)^T (w-v) = 0$$

$$\therefore [h(w)-h(v) \quad w-v] \cdot n = 0$$

$$\therefore n = \begin{bmatrix} 1 \\ -\nabla g(v) \end{bmatrix} \text{ is the normal vector to the tangent hyperplane}$$

Exercise 2.7

① $g(w) = w^2$

$\therefore g''(w) = 2 > 0$ for all w

$\therefore$ it is convex

② $g(w) = e^{w^2}$

$\therefore g''(w) = e^{w^2} \cdot (2w)^2 + 2e^{w^2} > 0$ for all w

$\therefore$ it is convex

③ $g(w) = \log(1+e^{w})$

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

$\therefore$ it is convex

④ $g(w) = -\log(w)$

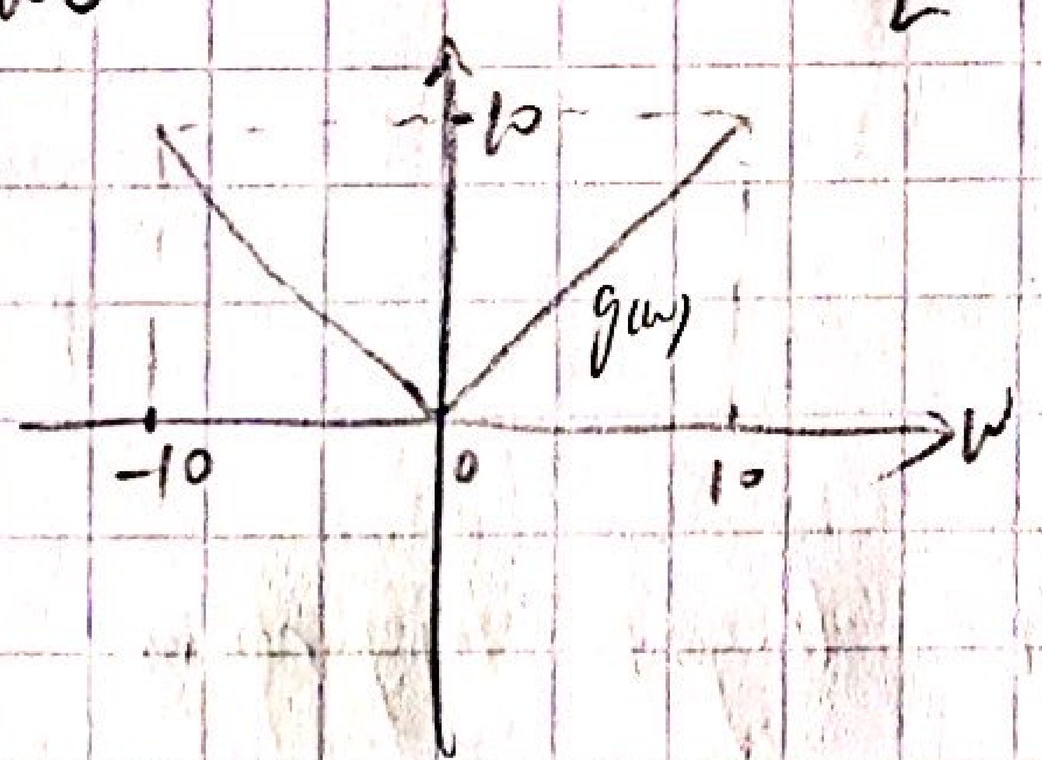
$\therefore g''(w) = \frac{1}{w^2} > 0$ for all w

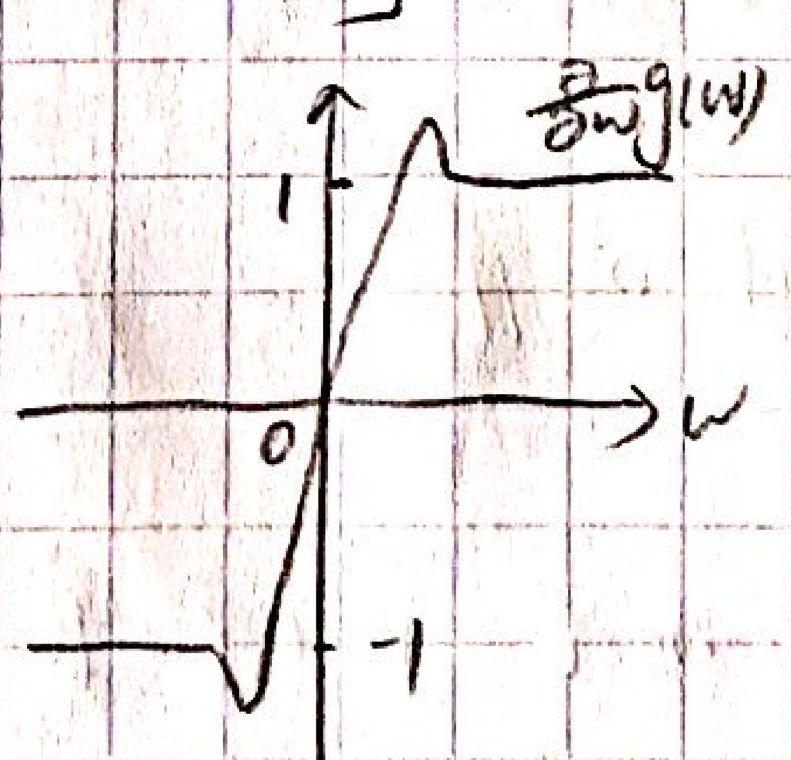
$\therefore$ it is convex

Exercise 2.8

a) $g(w) = w\tanh(w)$

$\therefore \frac{\partial}{\partial w} g(w) = \tanh(w) - w[\tanh(w^2) - 1]$





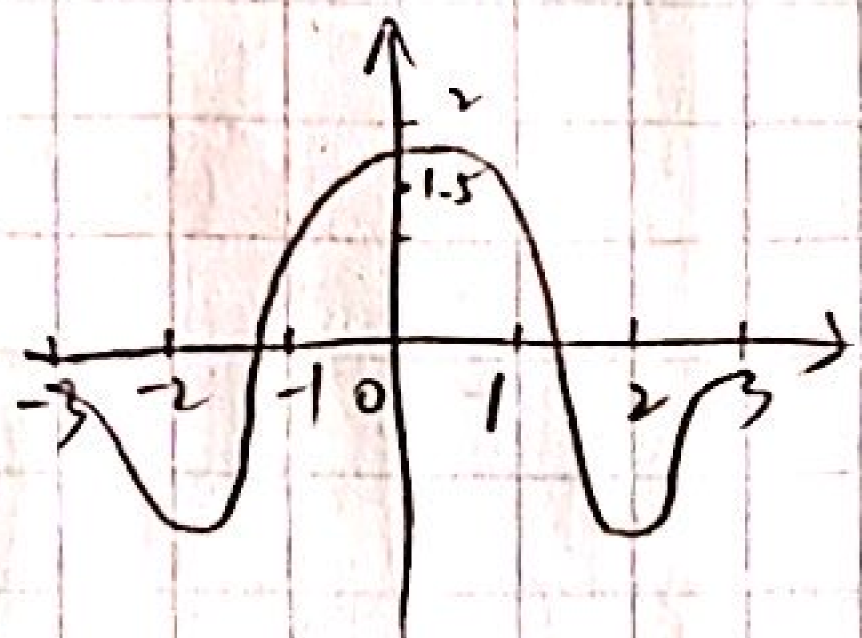
Using Matlab to draw these figures

$\therefore$ the stationary point of $g(w) = w\tanh(w)$ is $(0,0)$

EE475 HW1

Name: Zhisheng Lin
student #: 3226299

b) $\frac{\partial^2}{\partial w^2} g(w) = 2w\tanh(w)\left[\tanh(w^2) - 1\right] - 2\tanh(w^2) + 2$



Using Matlab to draw this figure

Since the $\frac{\partial^2}{\partial w^2} g(w)$ is not always bigger than 0,

g is non-convex.