Yuhang Peng NetID : ypa8256.

2.1

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

First: $g'(w) = \frac{\partial g(w)}{\partial w} = q w + r$

second: $g''(w) = \frac{\partial^2 g(w)}{\partial w^2} = q$

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

First: $g'(w) = [\sin(2\pi w^2)] \cdot 4\pi w + 2w$

second: $g''(w) = [\cos(2\pi w^2)] (4\pi w)^2 + [\sin(2\pi w^2)](4\pi) + 2$

$= 4\pi [4\pi w^2 \cos(2\pi w^2) + \sin(2\pi w^2)] + 2.$

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

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

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

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

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

2.2

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

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

Hessian:

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

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

Gradient: $\nabla_{\bar{w}} g(\bar{w}) = [\sin(2\pi\bar{w}^T\bar{w})](4\pi\bar{w}) + 2\bar{w}$

Hessian: $\nabla^2_{\bar{w}} g(\bar{w}) = [\cos(2\pi\bar{w}^T\bar{w})](4\pi\bar{w})^2 + [\sin(2\pi\bar{w}^T\bar{w})]4\pi + 2.$

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

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

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

2.5

for $h(\bar{w}) = g(\bar{v}) + \nabla g(\bar{v})^T(\bar{w} - \bar{v})$

$$h(\bar{w}) - g(\bar{v}) - \nabla^T g(\bar{v})\bar{w} + \nabla^T g(\bar{v})\bar{v} = 0$$

$$[1 \quad -\nabla^T g(\bar{v})] \begin{bmatrix} h(\bar{w}) \\ w_1 \\ \vdots \\ w_N \end{bmatrix} + \nabla^T g(\bar{v})\bar{v} - g(\bar{v}) = 0.$$

$$\underset{\sim}{n}^T \begin{bmatrix} h \\ w_1 \\ \vdots \\ w_N \end{bmatrix} + r = 0$$

2.7:

1) $g(w) = w^2$ $\quad g''(w) = 2 > 0$ $\quad \therefore$ convex

2) $g(w) = e^{w^2}$ $\quad g'(w) = 2e^{w^2} w,\; g''(w) = 2e^{w^2} + 4w^2 e^{w^2} > 0$ $\quad \therefore$ convex

3) $g(w) = \log(1+e^{w})$ $\quad g''(w) = \dfrac{2e^{w^2}(2w^2+e^{w^2}+1)}{(1+e^{w^2})^2} > 0$ $\quad \therefore$ convex

4) $g(w) = -\log(w)$ $\quad g'(w) = \dfrac{1}{\frac{1}{w}\ln 2}$ $\quad g''(w) = \dfrac{1}{\ln 2} > 0$ $\quad \therefore$ convex

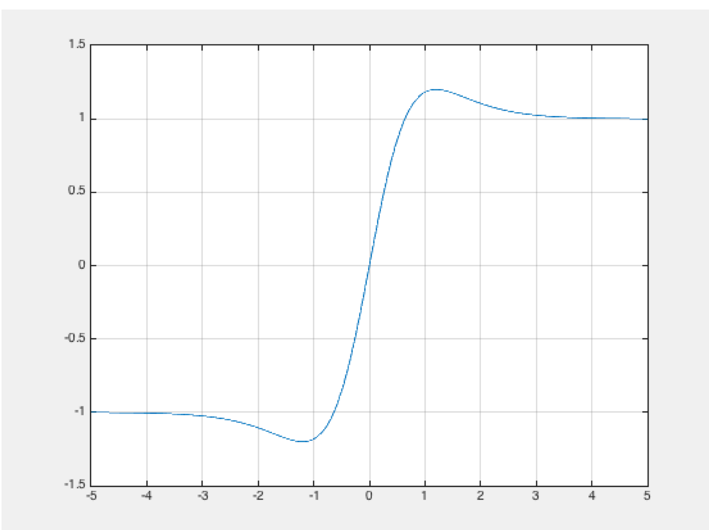
2.8

(1) $g(w) = w \cdot \tanh(w)$

$$\frac{dg(w)}{dw} = \tanh(w) - w \cdot \left(\left(\tanh(w)\right)^2 - 1\right)$$

After we that the first order equal to 0, then we can obtain that w = 0.

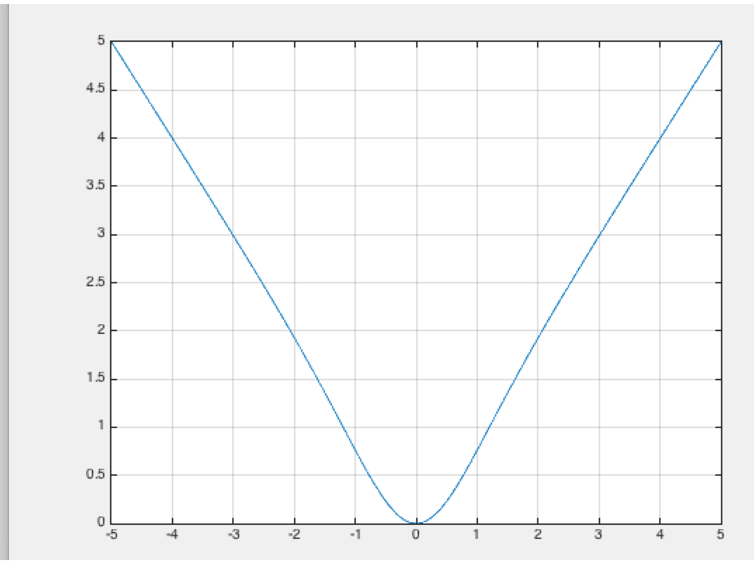
This figure show the $\frac{dg(w)}{dw}$, and we can find that w = 0 is the stationary point.




```
>> plot(x,tanh(x) - x.*(tanh(x).^2 - 1)),grid on
```

To check the answer, we plot the g(w), that shows the stationary point is the global minimal point.

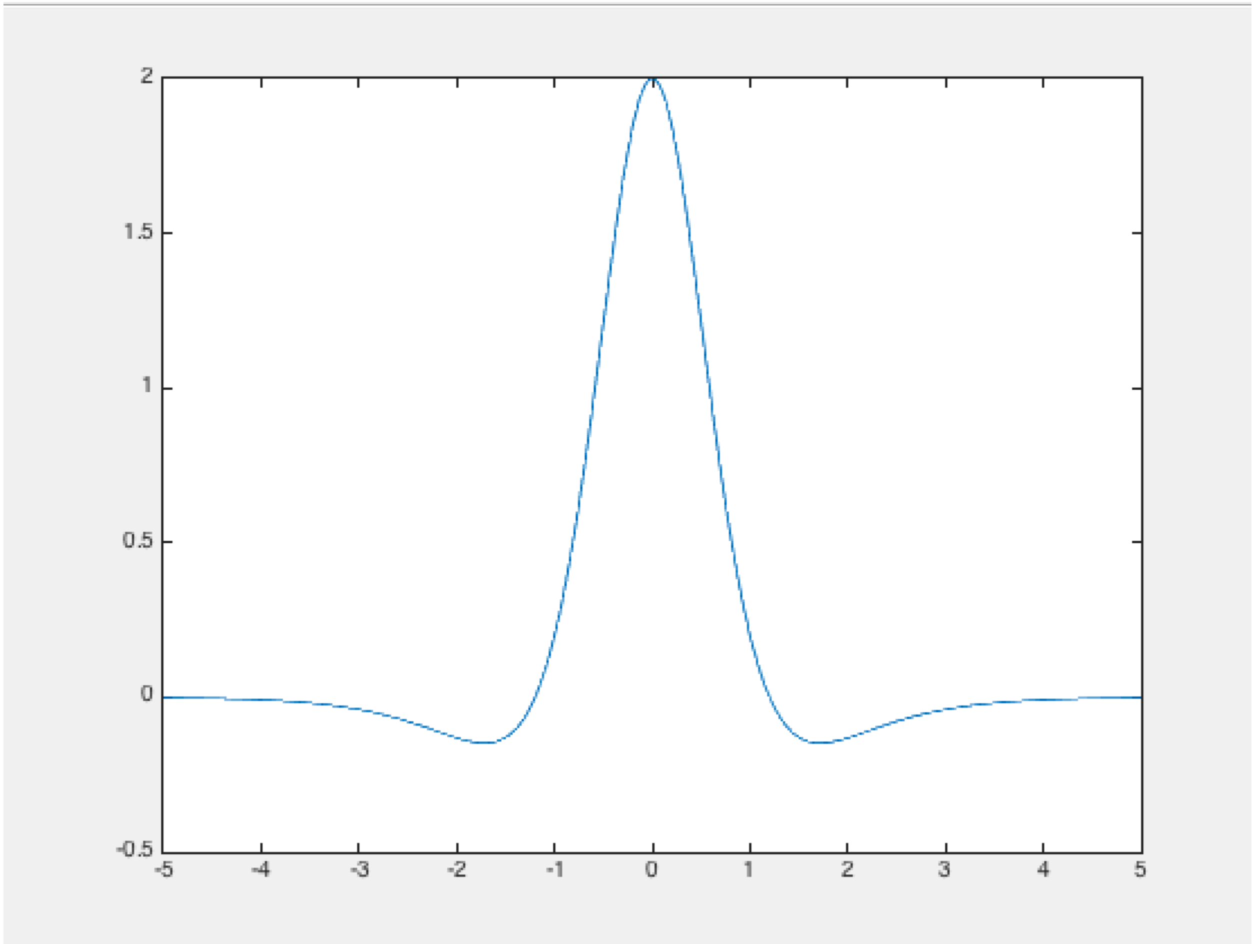
```
x = -5:0.01:5;
plot(x,x.*tanh(x)),grid on
```



(2)

After plot the second order of the g(w), g''(w) < 0 at some point, so it is non-convex

```
>> f = x.*tanh(x);
>> g = diff(f)

g =

tanh(x) - x*(tanh(x)^2 - 1)

>> k = diff(g)

k =

2*x*tanh(x)*(tanh(x)^2 - 1) - 2*tanh(x)^2 + 2

>> plot(x,2.*x.*tanh(x).*(tanh(x).^2 - 1) - 2.*tanh(x).^2 + 2)
```


2
1.5
1
0.5
0
-0.5
-5
-4
-3
-2
-1
0
1
2
3
4
5