3. [10 points] prove that the Hessian is positive semidefinite.

- A matrix is *positive semidefinite* if and only if $\mathbf{b}^T\boldsymbol{\mathcal{H}}\mathbf{b} > \vec{0} \; \forall \; \mathbf{b} \neq \vec{0}$, where $\mathbf{b}$ is a vector. $\boldsymbol{\mathcal{H}}(f(\mathbf{x}_i))$ can be re-written as, by using Einstein Summation Notation

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \frac{||\mathbf{w}||^2 y_i^2 e^{\varsigma_i}}{(e^{\varsigma_i}+1)^2} & \frac{y_i[e^{\varsigma_i}(\varsigma_i-1)-1]}{(e^{\varsigma_i}+1)^2} \\ \frac{y_i[e^{\varsigma_i}(\varsigma_i-1)-1]}{(e^{\varsigma_i}+1)^2} & \frac{||\mathbf{x}_i||^2 y_i^2 e^{\varsigma_i}}{(e^{\varsigma_i}+1)^2} + \frac{2}{\sigma^2} \end{bmatrix} \tag{56}$$

where $\varsigma_i = \mathbf{w}^T\mathbf{x}_i y_i$

We only care about if it is going to be negative, so without loss of generality we can remove all the norms, squares, and denominators since they are *always* positive. We can also generalize our vector $\mathbf{b}$ as $\mathbf{b} = (\alpha_1, \alpha_2)^T$, where α_1 and α_2 can be anything

$$\boldsymbol{\mathcal{H}}' = \begin{bmatrix} e^{\varsigma_i} & e^{\varsigma_i}(\varsigma_i - 1) - 1 \\ e^{\varsigma_i}(\varsigma_i - 1) - 1 & e^{\varsigma_i} + \frac{2}{\sigma^2} \end{bmatrix} \tag{57}$$

where $\mathbf{b}^T\boldsymbol{\mathcal{H}}'\mathbf{b}$ can be generalized with the following result

$$\mathbf{b}^T \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \mathbf{b} = x_1\alpha_1^2 + x_4\alpha_2^2 + \alpha_1\alpha_2(x_2 + x_3) \tag{58}$$

when applied to $\boldsymbol{\mathcal{H}}'$ results in

$$\mathbf{b}^T\boldsymbol{\mathcal{H}}'\mathbf{b} = \alpha_1^2 e^{\varsigma_i} + \alpha_2^2\left(e^{\varsigma_i} + \frac{2}{\sigma^2}\right) + 2\alpha_1\alpha_2\left(e^{\varsigma_i}(\varsigma_i - 1) - 1\right) \tag{59}$$

Where there's two instances that we need to prove. (1) If $\alpha_1, \alpha_2 > 0$ then $\alpha_1\alpha_2(x_2+x_3) > 0$ and (2) if $\alpha_1 > 0$ and $\alpha_2 < 0$, then $\alpha_1^2 x_1 + \alpha_2^2 x_2 > \alpha_1\alpha_2(x_2+x_3)$.

$$\alpha_1, \alpha_2 > 0 \Rightarrow \alpha_1\alpha_2(x_2 + x_3) > 0 \tag{60}$$

$$\alpha_1\alpha_2(x_2 + x_3) = 2\alpha_1\alpha_2\varsigma_i e^{\varsigma_i} - 2e^{\varsigma_i} - 2 > 0 \tag{61}$$

$$\underbrace{\varsigma_i e^{\varsigma_i} - e^{\varsigma_i} - 1}_{>\ 0 \text{ if } \varsigma_i e^{\varsigma_i} > e^{\varsigma_i} - 1}; \quad \varsigma_i \neq 0 \tag{62}$$

$$\Rightarrow \varsigma_i e^{\varsigma_i} > e^{\varsigma_i} - 1 \tag{63}$$

$$\Rightarrow \varsigma_i > 1 - e^{-\varsigma_i} \quad \square \tag{64}$$

where $\varsigma_i \geq 1 \; \forall \; i$. Now for the alternative case where if $\alpha_1 > 0$ and $\alpha_2 < 0$, that $\alpha_1^2 x_1 + \alpha_2^2 x_2 > \alpha_1\alpha_2(x_2 + x_3)$

$$\alpha_1^2 e^{\varsigma_i} + \alpha_2^2\left(e^{\varsigma_i} + \frac{2}{\sigma^2}\right) > \alpha_1\alpha_2\left[e^{\varsigma_i}(\varsigma_i - 1) - 1\right] \tag{65}$$

$$\left[\alpha_1^2 + \alpha_2^2 + \frac{2}{\sigma^2 e^{\varsigma_i}}\right] > \left[\alpha_1\alpha_2(\varsigma_i - 1) - e^{-\varsigma_i}\right] \tag{66}$$