Question 2.3 : $f_1(w) = ||w||^2 = w^T . w$

$f_2(\mathcal{E}) = \mathcal{E}$

Hessian matrix of $f_1(w)$ where $\bar{w} = \begin{bmatrix} w_0 \\ \vdots \\ w_n \end{bmatrix}$ should be positive semi-definite.

$$H \text{ of } f_1(w) = \begin{bmatrix} \frac{\partial^2 f_1}{\partial^2 w_0} & \frac{\partial^2 f_1}{\partial w_0 w_1} & \cdots & \frac{\partial^2 f_1}{\partial w_0 w_n} \\ \frac{\partial^2 f_1}{\partial w_1 w_0} & \frac{\partial^2 f_1}{\partial w_1^2} & \cdots & \frac{\partial f_1}{\partial w_1 w_n} \\ \vdots & & & \\ \frac{\partial^2 f_1}{\partial w_n w_0} & \frac{\partial^2 f_1}{\partial w_n w_1} & \cdots & \frac{\partial^2 f_1}{\partial w_n^2} \end{bmatrix}$$

$$\frac{\partial^2 f_1}{\partial w_0^2} = 2, \quad \frac{\partial f_1}{\partial w_0 w_1} = 0, \quad \ldots$$

$$\frac{\partial^2 f_1}{\partial w_i^2} = 2 \text{ for } i = 0, \ldots, n \quad \& \quad \frac{\partial^2 f_1}{\partial w_i w_j} = 0$$

for $i = 0, \ldots, n$ & $j = 0, \ldots, n$ $i \neq j$

$$\text{therefore } H \text{ of } f_1(w) = \begin{bmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 2 \end{bmatrix} = 2 . I$$

Testing whether H of $f_1(w)$ is semi definite positive

$$x = \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}$$

$$x^T . H . x = [x_0 \dots x_n] \underbrace{\begin{bmatrix} 2 & 0 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & 2 \end{bmatrix}}_{\substack{2I \\ \text{positive}}} \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} = 2 . \sum_{i=1}^{n} (x_i)^2 \geq 0$$

for all x.

Therefore H is a semi-definite which means $f_1(w)$ is convex.

Now considering $f_2(\varepsilon) = \varepsilon$.

if $f_2(\varepsilon)$ is convex on interval I, then for any $\varepsilon_1, \varepsilon_2 \in I$

$$\underbrace{f_2\left(\frac{\varepsilon_1 + \varepsilon_2}{2}\right)} \leq \frac{f_2(\varepsilon_1) + f_2(\varepsilon_2)}{2}$$

$$\frac{\varepsilon_1 + \varepsilon_2}{2} \leq \frac{\varepsilon_1 + \varepsilon_2}{2}$$ $\Rightarrow$ since this inequality holds

we can say $f_2(\varepsilon) = \varepsilon$ is convex function.

Since sum of convex functions will be convex too,

we can say $\sum_{i=1}^{n} \varepsilon_i$ is also convex function.

Therefore, $\underbrace{\|w\|^2}_{h(w)} + \underbrace{\sum_{i=1}^{n} \varepsilon_i}_{g(\varepsilon)}$ will be convex too since

we showed $h(w)$ & $g(\varepsilon)$ are convex functions & their sum

will be convex too.

Question 2.4 : $\min \|w\|^2 + \sum_{i=1}^{n} \varepsilon_i$

$$\text{s.t} \quad y_i w^T x_i \geq 1 - \varepsilon_i, \quad i = 1, \dots, n$$
$$\varepsilon_i \geq 0, \quad i = 1, \dots, n$$

we should write the above as :

$$\min f_0(w, \varepsilon)$$
$$\text{s.t} \quad g_i \leq 0, \quad i = 1, \dots, m$$
$$t_i \leq 0, \quad i = 1, \dots, m$$

$$y_i w^T x_i - 1 + \varepsilon_i \geq 0 \rightarrow \underbrace{-y_i w^T x_i + 1 - \varepsilon_i}_{g_i} \leq 0$$

$$\varepsilon_i \geq 0 \rightarrow \underbrace{-\varepsilon_i}_{t_i} \leq 0$$

$$f(x, \alpha, \beta) = \underbrace{\|w\|^2 + \sum_{i=1}^{n} \varepsilon_i}_{f_0(w,\varepsilon)} + \sum_{i=1}^{n} \alpha_i . (-y_i w^T x_i + 1 - \varepsilon_i) + \sum_{i=1}^{n} \beta_i . (-\varepsilon_i)$$

$$= \|w\|^2 + \sum_{i=1}^{n} \varepsilon_i - \sum_{i=1}^{n} \alpha_i . y_i w^T x_i + \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \alpha_i . \varepsilon_i - \sum_{i=1}^{n} \beta_i . \varepsilon_i$$

$$= \|w\|^2 + \sum_{i=1}^{n} (1 - \alpha_i - \beta_i) . \varepsilon_i + \sum_{i=1}^{n} \alpha_i - \sum_{i=1}^{n} \alpha_i y_i w^T x_i$$

Question 2.7:

Assume $\bar{x}$ is feasible,

$$\mathcal{L}(\bar{x}, \alpha, \beta) = f_0(\bar{x}) + \sum_{i=1}^{m} \alpha_i \cdot g_i(\bar{x}) + \sum_{i=1}^{p} \beta_i \cdot h_i(\bar{x})$$

we know that if $\bar{x}$ is feasible $g_i(\bar{x}) \leq 0$ & $h_i(\bar{x}) = 0$

hence Lagrangian reduces to:

$$\mathcal{L}(\bar{x}, \alpha, \beta) = f_0(\bar{x}) + \underbrace{\sum_{i=1}^{m} \underbrace{\alpha_i}_{\geq 0} \cdot \underbrace{g_i(\bar{x})}_{\leq 0}}_{\text{this term is non-positive here}}$$

it can be at maximum 0, or it's gonna be negative.

Therefore $\max(\mathcal{L}(\bar{x}, \alpha, \beta)) = f_0(\bar{x})$ where $\sum_{i=1}^{m} \alpha_i \cdot g_i = 0$

for any feasible $\bar{x}$.

Question 2.6 :

$$\max(f(x,\alpha,\beta)) = \max\left(f_0(x) + \sum_{i=1}^{m} \alpha_i . g_i(x) + \sum_{i=1}^{p} \beta_i . h_i(x)\right)$$

we know that $f_0(x)$ & $g_i(x)$ are convex

$\alpha_i . g_i(x)$ where $\alpha \geq 0$ will also be convex

and $\sum_{i=1}^{m} \alpha_i . g_i(x)$, therefore will be convex too.

for a feasible $\bar{x}$ the max of Lagrangian is written as:

$$\max(f(\bar{x},\alpha,\beta)) = \max\left(f_0(\bar{x}) + \sum_{i=1}^{m} \alpha_i . g_i(\bar{x})\right)$$

where $f_0(\bar{x}) + \sum_{i=1}^{m} \alpha_i . g_i(\bar{x})$ is convex.

from pointwise maximum property of convex functions, we can say $\max(f(\bar{x},\alpha,\beta))$ is also convex since $g = \max\{f_1, \ldots, f_n\}$ is convex for all convex f_i functions, Question 2.7 : assuming $\bar{x}$ is feasible

$$f(\bar{x}, \lambda, \nu) \leq f_0(\bar{x})$$

$$\min_x \{ f(\bar{x}, \lambda, \nu) \} \leq f(\bar{x}, \lambda, \nu) \leq \underbrace{f_0(\bar{x})}_{\max(f(\bar{x}, \lambda, \nu))}$$

we know for feasible $\bar{x}$

$$\Theta_P(\bar{x}) = f_0(\bar{x}) \quad \text{and} \quad \min_x f(x, \lambda, \nu) = \Theta_D(\lambda, \nu)$$

therefore

$$\rightarrow \Theta_D(\lambda, \nu) \leq f(\bar{x}, \lambda, \nu) \leq \Theta_P(\bar{x})$$

Question 2.7 :

$$\Theta_P(x^*) = \max_{\alpha \geq 0, \beta} f(x^*, \alpha, \beta) = f_0(x^*)$$

$$\Theta_D(\alpha^*, \beta^*) = \min_x f(x, \alpha^*, \beta^*)$$

for any α, β pairs & for $\bar{x}$ is feasible:

$$\min_x \{ f(\bar{x}, \alpha, \beta) \} \leq f(\bar{x}, \alpha, \beta) \leq \underbrace{f_0(\bar{x})}_{= \Theta_P(\bar{x})}$$

here $\min_x f(\bar{x}, \alpha^*, \beta^*) \leq \Theta_P(\bar{x})$

if x^* is optimal, it is also feasible. so we can insert x^* instead of $\bar{x}$.

$$\min_x f(x^*, \alpha^*, \beta^*) \leq \Theta_P(x^*)$$

$$= \Theta_D(\alpha^*, \beta^*) \leq \Theta_P(x^*)$$