Name : Md Faizal Karim
Roll No : 2022121004

1 (a) $f(x) = x\log x$

$\therefore f^*(y) = \sup_{x \in dom(f)} \{ yx - x\log x \}$

Now,

$dom(f) : x > 0, x \in \mathbb{R}$

we show $yx - x\log x$ is concave

Proof : let $f(x) = yx - x\log x$

$\Rightarrow f'(x) = y - (\frac{1}{\cancel{x}} + \log x)$

$\Rightarrow f''(x) = +\frac{1}{x} < 0$

$\therefore f''(x) < 0 \; \forall x \in dom(f)$

$\Rightarrow yx - \log x$ is concave.

Hence, to find the supremum of $yx - x\log x$, we can simply find its max value.

$\therefore f'(x) = y - (1 + \log x) = 0$

$\Rightarrow y = 1 + \log x \Rightarrow x = e^{y-1}$

$\therefore \cancel{f^*(y = 1+\log x)} = x(1 + \log x) - x\log x$

$= \cancel{x + x\log x - x\log x}$

$= \cancel{x}$

$\therefore f^*(y) = ye^{y-1} - e^{y-1}(y-1)$

$= e^{y-1}$

$\therefore \boxed{f^*(y) = e^{y-1}}$ , $dom(f^*) = \mathbb{R}$

(b) $f(x) = \frac{1}{x}$

$\therefore f^*(y) = \sup_{x \in dom(f)} \{ yx - \frac{1}{x} \}$

Now,

$dom(f) = x \neq 0$

We show $yx - \frac{1}{x}$ is concave for $x > 0$

Proof : Let $h(x) = yx - \frac{1}{x}$

$\Rightarrow h'(x) = y + \frac{1}{x^2}$

$\Rightarrow h''(x) = -\frac{2}{x^3}$

$\Rightarrow h''(x) < 0 \ \forall \ x > 0$

$\therefore yx - \frac{1}{x}$ is concave for $x > 0$

$\therefore$ Finding the max of $h$ : $h'(x) = 0$

$$\Rightarrow y + \frac{1}{x^2} = 0$$

$$\Rightarrow x = (-y)^{-\frac{1}{2}}$$

$$\therefore f^*(y) = y(-y)^{-\frac{1}{2}} - (-y)^{\frac{1}{2}}$$

$$= \frac{y - (-y)}{(-y)^{\frac{1}{2}}}$$

$$= \frac{2y}{(-y)^{\frac{1}{2}}} = +2(-y)^{\frac{1}{2}}, \quad y < 0$$

$$\therefore \boxed{f^*(y) = \cancel{2y}\ 2(-y)^{\frac{1}{2}}, \quad \text{dom}(f^*) : y < 0}$$

2. $f(x) = x\log x - (x+1)\log(x+1)$

$$\therefore D_f(p\|q) = \sum_x q(x)\, f\left(\frac{p(x)}{q(x)}\right)$$

$$= \sum_x \left[ p(x)\log p(x) - \left(\frac{p(x)+q(x)}{\cancel{q(x)}}\right) \log\left(\frac{p(x)}{q(x)} + 1\right)\right]$$

$$= \sum_x \left[ p(x)\log p(x) - p(x)\log q(x) - (p(x)+q(x))\log(p(x)+q(x)) + (p(x)+q(x))\log q(x)\right]$$

$$= \sum_x \left[ p(x)\log p(x) - \cancel{p(x)\log q(x)} - p(x)\log(p(x)+q(x)) - q(x)\log(p(x)+q(x)) + \cancel{p(x)\log q(x)} + q(x)\log q(x)\right]$$

$$= \sum_x \left[ p(x)\log\left(\frac{p(x)}{p(x)+q(x)}\right) + q(x)\log\left(\frac{q(x)}{p(x)+q(x)}\right)\right]$$

$$= \sum_x \left[ p(x)\log\left(\frac{2p(x)}{p(x)+q(x)}\right) + q(x)\log\left(\frac{2q(x)}{p(x)+q(x)}\right)\right] - \sum p(x)\log 2 - \sum q(x)\log 2$$

$$= \sum_x p(x)\log\left(\frac{p(x)}{\frac{p(x)+q(x)}{2}}\right) + \sum_x q(x)\log\left(\frac{q(x)}{\frac{p(x)+q(x)}{2}}\right) - 2\log 2$$

$$= 2\,\text{JSD}(p,q) - \log 4$$

$$\therefore \boxed{D_f(p\|q) = 2\,\text{JSD}(p,q) - \log 4.}$$

3. $f(x) = x\log x - (1+x)\log(1+x)$

$\therefore f^*(y) = \sup\limits_{x \in dom(f)} \{ yx - f(x) \}$

$= \sup\limits_{x \in dom(f)} \{ yx - x\log x + (x+1)\log(x+1) \}$

Now,

$dom(f): \; x > 0 \cap x > -1$

$= x > 0$

We show that $yx - x\log x + (x+1)\log(x+1)$ is concave

let $g(x) = yx - x\log x + (x+1)\log(x+1)$

$\Rightarrow g'(x) = y - (\log x + 1) + \log(x+1) + 1$

$= y - \log x + \log(x+1)$

$\Rightarrow g''(x) = -\frac{1}{x} + \frac{1}{x+1} = \frac{-x-1+x}{x(x+1)} = \frac{-1}{x(x+1)}$

$\Rightarrow g''(x) < 0 \;\forall\; x > 0 \;[\because dom(f): x > 0]$

$\therefore$ g(x) is concave.

$\therefore$ supremum of g(x) is its maximum value.

$\therefore g'(x) = y - \log x + \log(x+1) = 0$

$\Rightarrow \log\left(\frac{x+1}{x}\right) = y$

$\Rightarrow x = (-1+e^{-y})^{-1} = \cancel{\frac{1}{1+e^{y}}} \quad \frac{1}{e^{-y}-1}$

$\therefore \cancel{f^*(y) = \frac{y}{e^{-y}-1} - \frac{\log(e^{-y}-1)}{e^{-y}-1} + \frac{\log\left(\frac{e^{-y}}{e^{-y}+1}\right)}{e^{-y}+1} + \log\left(\frac{e^{-y}}{e^{-y}+1}\right)}$

$\cancel{= \frac{y}{e^{-y}-1} - \frac{\log(e^{-y}-1)}{e^{-y}-1} + \frac{-y}{e^{-y}+1} - \frac{\log(e^{-y}+1)}{e^{-y}+1} - y - \log(e^{-y}+1)}$

$\therefore f(x) = x\log x - (x+1)\log(x+1)$

$= x\log\left(\frac{x}{x+1}\right) - \log(x+1)$

$\Rightarrow f\left(x = \frac{1}{e^{-y}+1}\right) = \frac{1}{e^{-y}-1}\log\left(\frac{1}{e^{-y}}\right) - \log\left(\frac{e^{-y}}{e^{-y}-1}\right)$

$= \frac{y}{e^{-y}-1} - \log\left(\frac{e^{-y}}{e^{-y}-1}\right)$

$\therefore f^*(y) = \cancel{\frac{y}{e^{-y}-1}} - \cancel{\frac{y}{e^{-y}-1}} + \log\left(\frac{e^{-y}}{e^{-y}-1}\right)$

$= \log\left(\frac{1}{1-e^{-y}}\right) = -\log(1-e^{-y})$

$\therefore \boxed{f^*(y) = -\log(1-e^{-y})}$

4. A function is 1-Lipschitz if,

$$|f(x) - f(y)| \leq 1 \cdot \|x-y\| \quad , \forall \, x,y \in \text{dom}(f)$$

let $x = y+h$

$$\Rightarrow \left| f(y+h) - f(y) \right| \leq \| y+h-y \|_2$$

$$\Rightarrow | f(y+h) - f(y) | \leq \|h\|$$

$$\Rightarrow \frac{| f(y+h) - f(y) |}{\|h\|} \leq 1.$$

$$\Rightarrow \left| \frac{f(y+h) - f(y)}{h} \right| \leq 1.$$

$$\Rightarrow \lim_{h \to 0} \left| \frac{f(y+h) - f(y)}{h} \right| \leq \lim_{h \to 0} 1.$$

$$\Rightarrow | f'(y) | \leq 1.$$

$$\Rightarrow | \nabla f(x) | \leq 1 \quad \forall \, x \in \text{dom}(f)$$

$$\therefore \boxed{| \nabla f(x) | \leq 1 \; \forall \, x \in \text{dom}(f)}$$

7. $v^T (I - vv^T) v$

$= v^T ( v - v\underline{v^T v})$

$= v^T ( v - v) = 0 \quad [\because v^T v = 1]$

$\therefore$ $v$ and $(I - vv^T) v$ are orthogonal.

Next,

$(x^T v) v^T (I - vv^T) x$

$= v^T (x^T v) ( x - vv^T x)$

$= v^T (x^T v) x - v^T (x^T v) vv^T x$

let $x^T v$ be $k$

$= k v^T x - k v^T v v^T x$

$= k v^T x - k v^T x \quad [\because v^T v = 1]$

$= 0$

$\therefore$ $(x^T v) v$ and $(I - vv^T) x$ is orthogonal.

8.(a) $L(x,\nu) = x^Tx + \nu^T(Ax-b)$ , $x \in \mathbb{R}^n, A \in \mathbb{R}^{p\times n}, \nu \in \mathbb{R}^p, b \in \mathbb{R}^p$

(b) $g(\nu) = \inf_{x \in dom(f)} \left\{ x^Tx + \nu^T(Ax-b) \right\}$

$= \inf_{x \in \mathbb{R}^n} \left\{ x^Tx + \nu^T(Ax-b) \right\}$

~~(c) g'(v) =~~ We can see that the above expression is convex, hence, we can directly differentiate it and find the minima.

$\therefore$ let $f(x) = x^Tx + \nu^T(Ax-b)$

$\Rightarrow \nabla f(x) = 2x + A^T\nu = 0$

$\Rightarrow x = -\frac{1}{2}(A^T\nu)$

$\therefore f\left(x = -\frac{1}{2}(A^T\nu)\right) = +\frac{1}{2}(A^T\nu)^T\frac{(A^T\nu)}{2} + \nu^T\left(A \times -\frac{1}{2}(A^T\nu) - b\right)$

$= \frac{1}{4}\nu^TAA^T\nu + \frac{-\nu^T}{2}AA^T\nu - \nu^Tb$

$= -\frac{1}{4}\nu^TAA^T\nu - \nu^Tb$

$\therefore \boxed{g(\nu) = -\frac{1}{4}\nu^TAA^T\nu - \nu^Tb}$

(c) To check for concavity, we can show that $-\nabla^2 g$ is positive definite.

Also, $AA^T$ is a symmetric matrix —— ①

$\therefore \nabla g(\nu) = -\frac{1}{4} \times 2AA^T\nu - b$ $\left[ \because \nabla_x[x^TAx] = \begin{cases} 2Ax & \text{if } A \text{ is symmetric} \\ 0 & \text{if } A \text{ is skew-sym.} \end{cases} \right]$

$\Rightarrow -\nabla^2 g(\nu) = +\frac{1}{2}AA^T$

$\because$ $AA^T$ is a symmetric matrix $\Rightarrow$ it is positive semi-definite.

$\therefore -\nabla^2 g(\nu) > 0$

$\therefore$ $g(\nu)$ is concave.

(d) Slaters condition for this case:

→ There is no inequality to be satisfied

→ Primal problem is feasible

→ Case 1: $b \in R(A) \longrightarrow p^* = d^*$

Case 2: $b \notin R(A) \longrightarrow p^* = -\infty$ and $d^* = -\infty$

$\therefore$ Slaters condition is alway satisfied and Strong duality holds true. →(because of slater's condition)

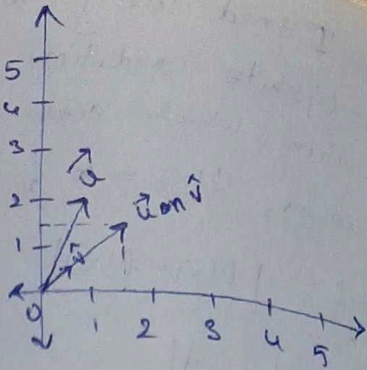
(e) Since, $p^* = d^*$, ~~then~~ hence the lower bound $d^*$ is tight to $p^*$.

7. $v = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^T$

let $u = [1, 2]^T$

$\therefore$ component of $u$ along $v$ = $\frac{u^T v}{\|v\|} \cdot \frac{v}{\|v\|}$

$$= \frac{[1\ 2]\begin{bmatrix}\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}}{1} \cdot \frac{\begin{bmatrix}\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\end{bmatrix}}{1}$$




$$= \frac{3}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]^T$$

$$= \left[\frac{3}{2}, \frac{3}{2}\right]^T$$

component of $u$ along $(I - vv^T)v$ = $u \cdot (I - vv^T)v$

= since $(I - vv^T)v$ is a null vector

$= [0, 0]^T$

6. Given L-Lipschitz function, $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$|f(x) - f(y)| \leq L \cdot \|x-y\|_2$ —— ①

Here, we chose distance metric $d() = \| \|_2$ (l-2 norm)

let $\epsilon > 0$ and $\delta = \frac{\epsilon}{L}$

$\therefore$ if $\|x-y\|_2 < \delta$

$\Rightarrow \|x-y\|_2 < \frac{\epsilon}{L}$

$\Rightarrow L \cdot \|x-y\|_2 < \epsilon$

$\Rightarrow |f(x) - f(y)| < \epsilon$ [from ①]

$\therefore |f(x) - f(y)| < \epsilon$

Hence from the given, if $\|x-y\|_2 < \delta$ then $|f(x) - f(y)| < \epsilon$, the function $f$ is continuous.

Hence, any L-Lipschitz continuous function is continuous.

Hence, proved.

5. I tried looking for a function $f: \mathbb{R}^n \to \mathbb{R}$ which follows 1-Lipschitz condition equality. But I couldn't guess any such function which goes from $\mathbb{R}^n \to \mathbb{R}$.

However, $f(x) = x$, $x \in \mathbb{R}$ satisfies the condition

Proof: $|f(x) - f(y)|$

$= |x - y|$

$= \|x - y\|_2$

Hence, $f(x) = x$, $x \in \mathbb{R}$ follows 1-Lipschitz equality.

a). $d(x,y) = \|x-y\|_2$, $x, y \in \mathbb{R}^n$

We first assume that $d(x,y)$ is convex and then prove it.

Proof: Using the condition of convexity,

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y), \quad x, y \in \mathbb{R}^n, \lambda \in [0,1]$$

$$\Rightarrow \|\lambda x + (1-\lambda)y\|_2 \le \lambda \|x\|_2 + (1-\lambda)\|y\|_2$$

$$\Rightarrow \|\lambda x + (1-\lambda)y\|_2^2 \le \lambda^2 \|x\|_2^2 + (1-\lambda)\|y\|_2^2 + 2\lambda(1-\lambda)\|x\|_2\|y\|_2$$

[squaring both sides]

a. $d(x,y) = \|x-y\|_2$, $x, y \in \mathbb{R}^n$

For simplicity we will denote $d()$ as $f(x) = \sqrt{x^T x}$, $x \in \mathbb{R}^n$

$$\therefore \nabla f(x) = \frac{1}{2\sqrt{x^T x}} \cdot 2x = \frac{x}{\|x\|_2} \quad [\because \nabla(x^T x) = 2x]$$

Now,

$$\nabla^2 f(x) = \nabla_x \left[ \frac{x}{\|x\|_2} \right]$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \begin{cases} \frac{1}{\|x\|_2} - \frac{x_i^2}{\|x\|_2^3} & i = j \\ 0 & i \neq j \end{cases}$$

$$\therefore \nabla^2 f(x) = \begin{bmatrix} c_1 & 0 & \cdots & 0 & 0 \\ 0 & c_2 & \cdots & 0 & 0 \\ \vdots & \vdots & c_3 & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & & c_n \end{bmatrix}_{n \times n}, \quad c_i = \frac{1}{\|x\|_2} - \frac{x_i^2}{\|x\|_2^3}$$

As we can see, $\nabla^2 f(x)$ has all diagonal elements non-zero and others are zero. Such a matrix always has positive eigenvalues.

$\therefore$, $\nabla^2 f(x)$ is positive definite.

$\therefore$, $f(x)$ is convex

$\Rightarrow$ L-2 norm is convex function.

Next,

Convexity is defined such that $\nabla^2 f(x) \succ 0$ and $dom(f)$ is a convex set. Above we did not show explicitly that $dom(f)$ is convex as $\mathbb{R}^n$ is convex is trivial.

But given $A \subset \mathbb{R}^n$ and $A$ is convex, $f(x)$ will still be convex since it meets both the conditions $\rightarrow$ $\nabla^2 f(x) \succ 0$ and $dom(f) = A$ is convex.

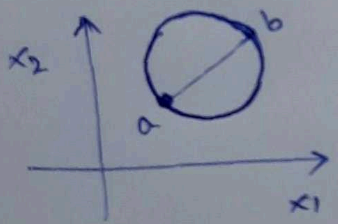
Hence, this proves the last statement of the question.

Next,

Let $A \subset \mathbb{R}^n$. $f(x)$ will not be convex in this $dom(f) = A$, if $A$ is non-convex. There are many examples of such subsets $\rightarrow$

One example is that of a <u>hollow circle</u>.

Suppose we consider the case of $A \subset \mathbb{R}^2$ for simplicity, we define $A$ as the points of a circle without the interior, only the boundary.




Here, if we connect any two points, it will include points which are not in the set. Hence, $A$ is non-convex.