(19) false take $n=2$

take a counter ex.

$$x = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad y = \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix}$$

$$\cancel{|x|^{t}} \quad \left| \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \times \lambda + (1-\lambda) \begin{bmatrix} 0 & 0 \\ 2 & 3 \end{bmatrix} \right|^{1/2} \leq \left| \begin{bmatrix} \lambda & \lambda \\ 2(1-\lambda) & \cancel{(1-\lambda)} \end{bmatrix} \right|^{1/2}$$

This is assumption that $f(x)$ is convex

$$\leq \lambda \left| \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix} \right|^{1/2} + \left| \begin{matrix} 0 & 0 \\ 2 & 1 \end{matrix} \right|^{1/2} (1-\lambda)$$

$$\left| \begin{bmatrix} \lambda & \lambda \\ 2(1-\lambda) & 3(1-\lambda) \end{bmatrix} \right|^{1/2} \leq 0$$

$$\cancel{\lambda^2 - \lambda} \; | \lambda - \lambda^2 - 2(1-\lambda)\lambda |^{1/2} \leq 0$$

$$| (1-\lambda)\lambda |^{1/2} \leq 0$$

for $\lambda \in (0,1)$

this inequality contradicts.

1 (b) False,

Proof :- Let's compute the hessian of function f first,

$\therefore f = \frac{x_1}{x_2}$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\Rightarrow \quad H = \begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & \frac{-2x_1}{x_2^3} \end{bmatrix}$$

As you can see that the above matrix is not positive-semi-definite

Reason being, if you try to find the eigen values,

$$|H - \lambda I| = \begin{vmatrix} -\lambda & -1/x_2^2 \\ -1/x_2^2 & \frac{-2x_1}{x_2^3} - \lambda \end{vmatrix} = \lambda\left(\frac{2x_1}{x_2^3} + \lambda\right) - \frac{1}{x_2^2} = 0$$

$$\Rightarrow \quad \lambda^2 + \frac{2\lambda x_1}{x_2^3} - \frac{1}{x_2^2} = 0$$

Observe that the coeff. of λ $\left(\frac{2x_1}{x_2^3} > 0\right)$

Hence ~~one~~ ~~atleast~~ one root has to be negative and therefore the eigenvalue would be negative $\forall\ x_1, x_2 \in \mathbb{R}_{++}$

$\therefore$ $H = \nabla^2 f$ is not positive semi-definite

and the function f is not convex.

[Clarification: Above we utilized the fact that All ~~positi~~ symmetric positive matrices have non-negative eigenvalues $\lambda \geq 0$]

① (c) $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$ on $\mathbb{R}^2_{++}$

~~False~~

True

Let's compute the Hessian of function $f(x_1, x_2) = \dfrac{1}{x_1 x_2}$ first.

$$H = \nabla^2 f(x) = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\Rightarrow H = \begin{bmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_1 x_2^3} \end{bmatrix}$$

When we

1c) ~~Tx~~ let $v = [v_1 \; v_2]^T$

then $$v^T H v = [v_1 \; v_2] \begin{bmatrix} \dfrac{2}{x_1^3 x_2} & \dfrac{1}{x_1^2 x_2^2} \\ \dfrac{1}{x_1^2 x_2^2} & \dfrac{2}{x_1 x_2^3} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$= [v_1 \; v_2] \begin{bmatrix} \dfrac{2v_1}{x_1^3 x_2} + \dfrac{v_2}{x_1^2 x_2^2} \\ \dfrac{v_1}{x_1^2 x_2^2} + \dfrac{2v_2}{x_1 x_2^3} \end{bmatrix}$$

$$= \frac{2v_1^2}{x_1^3 x_2} + \frac{2v_1 v_2}{x_1^2 x_2^2} + \frac{2v_2^2}{x_1 x_2^3}$$

$$= \frac{2}{x_1 x_2}\left[\Big(\underbrace{\frac{v_1}{x_1}}_{a}\Big)^2 + \Big(\frac{v_1}{x_1}\Big)\Big(\frac{v_2}{x_2}\Big) + \Big(\underbrace{\frac{v_2}{x_2}}_{b}\Big)^2 \right]$$

above expression is eq^n to

$$= \frac{2}{x_1 x_2}\left[\underbrace{a^2 + ab + b^2}_{\text{This quantity here is always } \geq 0} \right]$$

$\therefore$ H is a positive semi-definite matrix.

Q1 (d) Consider the function $g(x) = \|Ax - b\|$, for some $A \in R^{m \times n}$, $b \in R^m$

TRUE Consider the following Norm, for some $1 > \lambda > 0$

$$\| A(\lambda x + (1-\lambda)y) \mp b \|$$

{ Clarification: (−) sign before b }

$$= \| \lambda Ax + (1-\lambda)Ay - (\lambda + (1-\lambda))b \|$$

$$= \| \lambda(Ax - b) + (1-\lambda)(Ay - b) \|$$

Using the norm inequality,

$$\Rightarrow \quad \| \lambda(Ax-b) + (1-\lambda)(Ay-b) \|$$

$$\leq \lambda \|Ax - b\| + (1-\lambda)\|Ay - b\|$$

Hence the function, $g(x)$ is a convex function.

Originally, we had $f(x) = \max(g_1(x), g_2(x) \ldots, g_K(x))$

where $g_i(x) = \|A^{(i)}x - b^{(i)}\|$, for where $A^{(i)} \in R^{m \times n}$, $b^{(i)} \in R^{(i)}$

$\forall\ i = 1, 2, \ldots K.$

Since, it is known that pointwise maximum of convex function is also a Convex function (As discussed in the class), therefore $f(x)$ is also convex.

① ~~(d)~~ (e) Given $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$

False

~~Applying~~ Now, $f'(x) = \frac{d}{dx}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \right)$

Applying Newton Leibniz formula,

$$f'(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\Rightarrow f''(x) = \frac{1}{\sqrt{2\pi}} -x e^{-x^2/2}$$

As it can be seen above that the double derivation of the function $f(x)$,

$$f''(x) \geq 0 \quad \forall\, x \leq 0$$

else $f''(x) < 0 \quad \forall\, x > 0$.

Hence, the function $f(x)$ is not convex.

(2) Given that $d_1 > 0$ and $\lambda_i \leq 0 \;\forall\, i = 2, \ldots, k$. Also $\sum_{i=1}^{k} d_i = 1$.
We are expected to prove the following inequality,

$$f\left(\sum_{i=1}^{k} d_i x_i\right) \geq \sum_{i=1}^{k} d_i f(x_i)$$

We use the expanded version of convexity inequality (can be proven easily using induction) :-

Claim :- for $\alpha_i \geq 0$ with $\sum_{i=1}^{k} \alpha_i = 1$ and convex f.

$$\boxed{\sum_{i=1}^{k} \alpha_i f(y_i) \geq f\left(\sum_{i=1}^{n} \alpha_i y_i\right)}$$

Proof :- We use induction to prove this inequality

Base Case : for $k=2$, we get the usual convexity relation which is obviously true.

Suppose, that the relation holds true $\forall\, k < n$. We prove it for $k=n$.
Consider the following inequality.

$$\sum_{i=1}^{n} \alpha_i f(x_i) = \alpha_1 f(x_1) + (1-\alpha_1) \sum_{i=2}^{n} \frac{\alpha_i}{(1-\alpha_1)} f(x_i)$$

$$\geq \alpha_1 f(x_1) + (1-\alpha_1) f\left(\sum_{i=2}^{n} \frac{\alpha_i}{1-\alpha_i} x_i\right)$$

$$\geq f\left(\alpha_1 x_1 + (1-\alpha_1)\left(\sum_{i=2}^{n} \frac{\alpha_i}{1-\alpha_i} x_i\right)\right)$$

$$= f\left(\sum_{i=1}^{n} \alpha_i x_i\right)$$

Above, we first used the induction hypothesis for $k = n-1$ (as $\sum_{i=2}^{k} \frac{\alpha_i}{1-\alpha_i} = 1$) and then the convexity inequality with 2 terms. This proves the result expected and hence the claim is proved.

Now, we will use the above result to prove this inequality.
Let $\alpha_i = \frac{1}{d_i}$, $\alpha_i = \frac{-d_i}{d_1}$ $\forall i = 2, \ldots, k$. Also put $y_1 = \sum_{i=1}^{k} d_i x_i$, $y_i = x_i$ $\forall\, i = 2, \ldots, k$. Clearly $\alpha_i \geq 0$. $\forall i = 2, \ldots, k$ and $\sum_{i=1}^{k} \alpha_i = 1$

We get the following :-

$$\frac{1}{\lambda_1} f\left(\sum_{i=1}^{k} \lambda_i x_i\right) + \sum_{i=2}^{k} \frac{-\lambda_i}{\lambda_1} f(x_i) \geqslant f\left(\frac{1}{\lambda_1}\left(\sum_{i=1}^{k} \lambda_i x_i\right) - \sum_{i=2}^{k} \frac{\lambda_i}{\lambda_1} x_i\right)$$

$$= f(x_i)$$

$$\Rightarrow \boxed{f\left(\sum_{i=1}^{k} \lambda_i x_i\right) \geqslant \sum_{i=1}^{k} \lambda_i f(x_i)}$$

Hence, we have proven the result.

Q3 (a) (i) We need to prove that for $t \in (0,1)$

$$\boxed{t^t (1-t)^{1-t} \geq \frac{1}{2}}$$

Let $f(t) = t^t (1-t)^{1-t}$ and $g(t) = \log f(t) = t \log t + (1-t)\log(1-t)$

$\because$ $t \in (0,1)$, $\therefore$ $g(t)$ is differentiable.

$$\frac{\partial g(t)}{\partial t} = \log t - \log(1-t) + \cancel{t} \times \frac{1}{\cancel{t}} - \cancel{(1-t)}\frac{1}{\cancel{(1-t)}}$$

$$\frac{\partial g(t)}{\partial t} = \log t - \log(1-t) + 1 - 1 = \log t - \log(1-t)$$

$$\Rightarrow \frac{\partial^2 g(t)}{\partial t^2} = \frac{1}{t} + \frac{1}{1-t} = \frac{1}{t(1-t)} > 0, \text{ as } t \in (0,1)$$

Hence the function $g(t)$ is a convex function.

Therefore, the point where $\frac{\partial g(t)}{\partial t} = 0$, gives minima.

$$\Rightarrow \log t = \log(1-t)$$

$$\Rightarrow \boxed{t = 1/2} \rightarrow \text{Minima}$$

$$\Rightarrow g(t) \geq g(1/2)$$

$$\Rightarrow g(t) = \log f(t) \geq \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\left(\frac{1}{2}\right) = \log\left(\frac{1}{2}\right)$$

$$\Rightarrow f(t) \geq \frac{1}{2}$$

Hence the Result asked in the problem has been proven.

Q3 a (ii) Consider the ratios,

for $t \leq 1/2$
$$\frac{a^{1-t} b^{t}}{a^{1/2} b^{1/2}} = \frac{a^{\frac{1}{2}-t}}{b^{1/2-t}} = \left(\frac{a}{b}\right)^{\frac{1}{2}-t} \leq 1$$
$\{ \because a<b \}$
$$\therefore \quad a^{1-t} b^{t} \leq \sqrt{ab}$$

for $t > 1/2$,
$$\frac{a^{t} b^{1-t}}{a^{1/2} b^{1/2}} = \frac{a^{t-1/2}}{b^{t-1/2}} = \left(\frac{a}{b}\right)^{t-1/2} \leq 1$$
$\{ \because a<b$ and $t > 1/2 \}$
$$\Rightarrow \quad a^{t} b^{1-t} \leq \sqrt{ab} \quad , t > 1/2$$

For the second part of this question, we will use AM-GM inequality.
$$\Rightarrow \quad \frac{a^{1-t} b^{t} + a^{t} b^{1-t}}{2} \geq \sqrt{(a^{1-t} b^{t})(a^{t} b^{1-t})}$$
$$\Rightarrow \quad \frac{a^{1-t} b^{t} + a^{t} b^{1-t}}{2} \geq \sqrt{ab} \quad , \text{Hence Proved.} \rightarrow (1)$$

Now, consider the expression,
$$\frac{a^{1-t} b^{t} + a^{t} b^{1-t}}{2} = \frac{(a^{1-t} - b^{1-t}) b^{t} + a^{t}(-a^{1-t} + b^{1-t}) + b + a}{2}$$
$$\Rightarrow \quad \frac{a^{1-t} b^{t} + a^{t} b^{1-t}}{2} = \frac{(a^{1-t} - b^{1-t}) b^{t} - a^{t}(a^{1-t} - b^{1-t}) + b + a}{2}$$
$$= \frac{a+b - (b^{t} - a^{t})(b^{1-t} - a^{1-t})}{2}$$
$\because \quad b > a$

$\therefore$ the quantity $(b^{t} - a^{t})(b^{1-t} - a^{1-t}) > 0$
$$\therefore \quad \boxed{\frac{a^{1-t} b^{t} + a^{t} b^{1-t}}{2} \leq \frac{a+b}{2}} \rightarrow (2)$$

Thereby, using the results (1) and (2) we can write

$$\boxed{\sqrt{ab} \le \frac{a^{1-t}b^{t} + a^{t}b^{1-t}}{2} \le \frac{a+b}{2}}$$

Hence Proved.

③ (b) $\because$ $f(x)$ is log-convex and positive function,

we have $f(\lambda x + (1-\lambda)y) \le f^{\lambda}(x) f^{1-\lambda}(y)$

Also on apply ~~A~~ weighted AM-GM inequality

$$\boxed{f^{\lambda}(x) f^{1-\lambda}(y) \le \lambda f(x) + (1-\lambda) f(y)}$$

$$\Rightarrow \boxed{f(\lambda x + (1-\lambda)y) \le f^{\lambda}(x) f^{1-\lambda}(y) \le \lambda f(x) + (1-\lambda) f(y)}$$

It implies that f is also convex.

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

$$\Rightarrow \int_0^1 f(\lambda x + (1-\lambda)y)\, d\lambda \le \int_0^1 \lambda f(x)\, d\lambda + \int_0^1 (1-\lambda) f(y)\, d\lambda$$

$\Rightarrow$ $\left\{ \text{Take } z = \lambda x + (1-\lambda) y \quad \frac{dz}{x-y} = d\lambda \right\}$

$$\Rightarrow \frac{1}{x-y} \int_y^x f(z)\, dz \le \frac{f(x) + f(y)}{2}$$

Put $x = b$ and $y = a$ Here

we get,

$$\boxed{\frac{1}{b-a} \int_a^b f(x)\, dx \le \frac{f(a) + f(b)}{2}} \rightarrow ①$$

Now, consider the integral $\int_a^b f(x)\, dx$

$$\Rightarrow \int_a^b f(x)\, dx = \int_a^b \frac{f(x) + f(a+b-x)}{2}\, dx \rightarrow ②$$

Since f is also convex,

take $\lambda = 1/2$ and $\alpha = x$, $\beta = a+b-x$

$$\therefore f\left(\frac{a+b}{2}\right) \le \frac{1}{2} f(x) + \frac{1}{2} f(a+b-x).$$ On taking integral a to b

$$\Rightarrow \int_a^b f\left(\frac{a+b}{2}\right) dx \le \int_a^b f(x)\, dx \quad \{\text{using } ②\}.$$

$$(b-a) \times f\left(\frac{a+b}{2}\right) \leq \int_a^b f(x)\,dx$$

$$\Rightarrow \quad f\left(\frac{a+b}{2}\right) \leq \frac{1}{(b-a)} \int_a^b f(x)\,dx \longrightarrow (3)$$

Therefore, using the results from (1) and (3) we have,

$$\boxed{f\left(\frac{a+b}{2}\right) \leq \frac{1}{(b-a)} \int_a^b f(x)\,dx \leq \frac{f(a)+f(b)}{2}}$$

Hence Proved,

(c) In part 3(a)(ii), we proved that

$$\sqrt{ab} \le \frac{a^{1-t}b^{t} + a^{t}b^{1-t}}{2} \le \frac{a+b}{2}$$

for $0 < a < b$

On integrating wrt t from 0 to 1.

$$\int_0^1 \sqrt{ab}\,dt \le \frac{\int_0^1 a^{1-t}b^{t}\,dt + \int_0^1 a^{t}b^{1-t}\,dt}{2} \le \int_0^1 \frac{a+b}{2}\,dt$$

$$\Rightarrow \sqrt{ab} \le \frac{a\int_0^1 \left(\frac{b}{a}\right)^t dt + b\int_0^1 \left(\frac{a}{b}\right)^t dt}{2} \le \frac{a+b}{2}$$

$$\Rightarrow \sqrt{ab} \le \frac{a\,\frac{\frac{b}{a}-1}{\ln(b/a)} + b\,\frac{\frac{a}{b}-1}{\ln(a/b)}}{2} \le \frac{a+b}{2}$$

$$\Rightarrow \sqrt{ab} \le \frac{\frac{b-a}{\ln(b/a)} + \frac{a-b}{\ln(a/b)}}{2} \le \frac{a+b}{2}$$

$$\Rightarrow \sqrt{ab} \le \frac{b-a}{\ln(b/a)} \le \frac{(a+b)}{2}$$

$$\Rightarrow \sqrt{ab} \le \frac{a-b}{\ln(a/b)} \le \frac{a+b}{2}$$

$$\Rightarrow \boxed{\sqrt{ab} \le L(a,b) \le \frac{a+b}{2}}$$ Hence proved.

Given $\phi(a,b) = \sqrt{f\left(\frac{3a+b}{4}\right) f\left(\frac{a+3b}{4}\right)}$

$\because$ f is log-convex choose $x = \frac{3a+b}{4}$, $y = \frac{a+3b}{4}$, $\lambda = \frac{1}{2}$

$\therefore$ Using $f(\lambda x + (1-\lambda) y) \leq (f(x)^{\lambda} f^{1-\lambda}(y))$

we have, $f\left(\frac{a+b}{2}\right) \leq \phi(a,b) \longrightarrow (1)$

Also using the previous result, $L(a,b) \leq \frac{a+b}{2}$

Use $a \longrightarrow f(a)$, $b \longrightarrow f(b)$

$\therefore$ $L(f(a), f(b)) \leq \frac{f(a) + f(b)}{2} \longrightarrow (2)$

(4) Given $|f'(x)| \geq c$ (Assuming zeta to be c)

let $t = f(x)$

$\Rightarrow dt = f'(x)\,dx$

$\Rightarrow dx = \dfrac{dt}{f'(x)}$

$\because \dfrac{1}{|f'(x)|} \leq \dfrac{1}{c}$

$\because c > 0$ therefore $f'(x)$ can never be 0, therefore it is monotonic in the range $[a, b]$

It's fair to Assuming that $f'(x) > 0$

$\therefore$ We can write $\left| \displaystyle\int_{f(a)}^{f(b)} \frac{e^{it}\,dt}{f'(x)} \right| < \left| \displaystyle\int_{f(a)}^{f(b)} \frac{e^{it}\,dt}{c} \right|$

$$\Rightarrow \left| \int_a^b e^{if(x)}\,dx \right| < \left| \int_{f(a)}^{f(b)} \frac{e^{it}\,dt}{c} \right| \longrightarrow (1)$$

$e^{i\theta} = \cos\theta + i\sin\theta$

$$\Rightarrow \frac{1}{c}\left| \int_{f(a)}^{f(b)} (\cos t + i\sin t)\,dt \right| = \frac{1}{c}\Big| \big(\cos(f(b)) - \cos(f(a))\big) + i\big(\sin(f(b)) - \sin(f(a))\big) \Big|$$

$$\Rightarrow \quad = \frac{1}{c}\left((\cos(m) - \cos(n))^2 + (\sin(m) - \sin(n))^2 \right)^{1/2}$$

$$= \frac{1}{c}\left(2 - 2\cos(m-n) \right)^{1/2}$$

for $\boxed{m - n = \frac{3\pi}{4}}$

$$< \frac{1}{c}(2+2)^{1/2} = \frac{2}{c}$$

$$\Rightarrow \boxed{\left| \int_a^b e^{if(x)}\,dx \right| < \frac{2}{c}}$$

Q5 Given : $\underbrace{L_i(\theta_i)}_{\text{Loss at the ith iteration}} = \underbrace{E_{(s,a,r,s')\sim U(D)}}_{\text{Represents experience is being drawn from Uniform Distribution.}} \left[\left(r + \gamma \max_{a'} Q(s', a'; \theta_i^-) - Q(s,a,\theta_i) \right)^2 \right]$

For the sake of simplicity in notation, lets remove subscript i above. Now we need to find $\nabla_\theta L(\theta)$. Before we proceed, observe that the $e = (s,a,r,s')$ follows a uniform distribution, (let the probability density function be denoted as $P(e)$).

And we need to find gradient w.r.t θ which is independent of e, therefore, we can write,

$$\nabla_\theta L(\theta) = E\left[\nabla_\theta \left(r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s,a,\theta) \right)^2 \right]$$

$$\Rightarrow \nabla_\theta L(\theta) = E\left[2\left(r + \gamma \max_{a'} Q(s', a', \theta^-) - Q(s,a,\theta) \right) \left(-\nabla_\theta Q(s,a,\theta) \right) \right]$$

Here r, $Q(s', a', \theta^-)$, γ are independent of θ $\therefore$ gradient of these quantities will be 0

$$\Rightarrow \boxed{\nabla_\theta L(\theta) = -2E\left[\left(r + \gamma \max_{a'} Q(s', a', \theta^-) - Q(s,a,\theta) \right) \left(\nabla_\theta Q(s,a,\theta) \right) \right]}$$

(6a) Claim : for non-negative points x_i, $\forall$ $\{i = 1, 2 \ldots n\}$ and ~~non~~ positive numbers $p_1, p_2, \ldots, p_n$ s.t. $\boxed{\sum p_i = 1}$

we have

$$\boxed{\prod_{i=1}^{n} x_i^{p_i} \leq \sum_{i=1}^{n} p_i x_i}$$

{ Also known as Weighted AM–GM Inequality

Proof :→ We know that the function $f(x) = \ln x$ is a concave function, and for concave functions we have the inequality

$$\boxed{f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \geq \sum_{i=1}^{n} \alpha_i f(x_i)}$$

$\therefore$ for $\alpha_i = p_i$ and $f(x) = \ln(x) = \ln(x)$

$$\ln\left(\sum_{i=1}^{n} \alpha_i x_i\right) \geq \sum_{i=1}^{n} \alpha_i \ln(x_i)$$

$$\Rightarrow \ln\left(\sum_{i=1}^{n} p_i x_i\right) \geq \sum_{i=1}^{n} \ln\left(x_i^{p_i}\right)$$

$$\Rightarrow \ln\left(\sum_{i=1}^{n} p_i x_i\right) \geq \ln\left(\prod_{i=1}^{n} x_i^{p_i}\right)$$

$$\Rightarrow \boxed{\sum_{i=1}^{n} p_i x_i \geq \prod_{i=1}^{n} x_i^{p_i}} \longrightarrow (1)$$

Hence Our claim is proved. This proves first part of ~~are~~ our ineq.

Now, Observe that $\boxed{\sum_{i=1}^{n} \frac{1-p_i}{n-1} = 1}$

$\therefore$ We can apply ~~AM~~ weighted AM–GM again for this time making the replacement

$$p_i \longrightarrow \left(\frac{1-p_i}{n-1}\right)$$

$$\Rightarrow \prod_{i=1}^{n} (x_i)^{\frac{1-p_i}{n-1}} \le \sum_{i=1}^{n} \frac{(1-p_i)}{n-1} x_i$$

$$\Rightarrow (n-1) \prod_{i=1}^{n} (x_i)^{\frac{1-p_i}{n-1}} \le \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} p_i x_i$$

$$\Rightarrow \boxed{\sum_{i=1}^{n} p_i x_i \le \sum_{i=1}^{n} x_i - (n-1) \prod_{i=1}^{n} (x_i)^{\frac{1-p_i}{n-1}}} \rightarrow (2)$$

Hence, combining the results (1) and (2), we get,

$$\boxed{\prod_{i=1}^{n} x_i^{p_i} \le \sum_{i=1}^{n} p_i x_i \le \sum_{i=1}^{n} x_i - (n-1) \prod_{i=1}^{n} (x_i)^{\frac{1-p_i}{n-1}}}$$

Hence Proved

(6) (b) Since f is a non-decreasing function, and also we found in the previous part that

$$\boxed{\prod_{i=1}^{n} x_i^{p_i} \le \sum_{i=1}^{n} p_i x_i}$$

therefore

$$f\left(\prod_{i=1}^{n} x_i^{p_i}\right) \le f\left(\sum_{i=1}^{n} p_i x_i\right)$$

Also, by the definition of convex functions, we have

$$f\left(\sum_{i=1}^{n} p_i x_i\right) \le \sum_{i=1}^{n} p_i f(x_i) \rightarrow (6) \quad \left\{ \sum_{i=1}^{n} p_i = 1 \text{ is given} \right\}.$$

Since, $\boxed{\sum_{i=1}^{n} \frac{1-p_i}{n-1} = 1}$

$\therefore$ By definition of convexity for f, we have

$$f\left(\sum_{i=1}^{n} \frac{1-p_i}{n-1} x_i\right) \le \sum_{i=1}^{n} \frac{(1-p_i)}{n-1} f(x_i) \rightarrow (3)$$

$$\because \prod_{i=1}^{n} (x_i)^{\frac{1-p_i}{n-1}} \le \sum_{i=1}^{n} \frac{(1-p_i)}{n-1} x_i$$

Also f is non-decreasing function, therefore,

$$f\left(\prod_{i=1}^{n}(x_i)^{\frac{1-p_i}{n-1}}\right) \le f\left(\sum_{i=1}^{n}\frac{(1-p_i)x_i}{n-1}\right) \rightarrow (4)$$

Using Results (3) and (4), we have.

$$f\left(\prod_{i=1}^{n}(x_i)^{\frac{1-p_i}{n-1}}\right) \le \sum_{i=1}^{n}\frac{(1-p_i)}{n-1}f(x_i)$$

$$\Rightarrow (n-1)\, f\left(\prod_{i=1}^{n}(x_i)^{\frac{1-p_i}{n-1}}\right) \le \sum_{i=1}^{n}f(x_i) - \sum_{i=1}^{n}p_i f(x_i)$$

$$\Rightarrow \sum_{i=1}^{n}p_i f(x_i) \le \sum_{i=1}^{n}f(x_i) - (n-1)\, f\left(\prod_{i=1}^{n}(x_i)^{\frac{1-p_i}{n-1}}\right)$$

Using the results (5) and (6), we can write. $\rightarrow (5)$

$$\boxed{f\left(\sum_{i=1}^{n}p_i x_i\right) \le \sum_{i=1}^{n}p_i f(x_i) \le \sum_{i=1}^{n}f(x_i) - (n-1)\, f\left(\prod_{i=1}^{n}(x_i)^{\frac{1-p_i}{n-1}}\right)}$$

Hence Proved,

(7) (a) We need to prove that the two definitions are equivalent. We can do that by showing for $L>0$, $\|\nabla f(x)-\nabla f(y)\| \leq L\|x-y\|$ implies $|f(x)-f(y)-\langle \nabla f(x), y-x\rangle| \leq \frac{L}{2}\|y-x\|_2^2$

Since, it is known that $G(b)-G(a) = \int_a^b g(\theta)\,d\theta$

Next, let $g(\tau) = \langle \nabla f(x+\tau(y-x)), y-x\rangle$ be a function in τ

Now, consider the integral of $g(\tau)$ from 0 and 1.

$$G(1)-G(0) = \int_0^1 \langle \nabla f(x+\tau(y-x)), y-x\rangle\, d\tau$$

Observe that, if we put $z = x+\tau(y-x)$

$$\Rightarrow dz = (y-x)\,d\tau$$

$$\Rightarrow \boxed{G(1)-G(0) = \int_x^y \nabla f(z)^T dz} = f(y)-f(x)$$

$$\therefore f(y)-f(x) = \int_0^1 \langle \nabla f(x+\tau(y-x)), y-x\rangle\, d\tau$$

$$\Rightarrow f(y)-f(x) = \int_0^1 \langle \nabla f(x+\tau(y-x)) - \nabla f(x) + \nabla f(x), y-x\rangle\, d\tau$$

$$\Rightarrow f(y)-f(x) = \langle \nabla f(x), y-x\rangle + \int_0^1 \langle \nabla f(x+\tau(y-x)) - \nabla f(x), y-x\rangle\, d\tau$$

$$\Rightarrow T = |f(y)-f(x)-\langle \nabla f(x), y-x\rangle| = \left|\int_0^1 \langle \nabla f(x+\tau(y-x)) - \nabla f(x), y-x\rangle\, d\tau\right|$$

$$\Rightarrow T \leq \int_0^1 |\langle \nabla f(x+\tau(y-x)) - \nabla f(x), y-x\rangle|\, d\tau$$

Using Cauchy-Schwartz inequality

$$\Rightarrow T \leq \int_0^1 \underbrace{\|\nabla f(x+\tau(y-x)) - \nabla f(x)\|}_{\text{We use Lipschitz gradient inequality here.}} \cdot \|y-x\|\, d\tau$$

$$\Rightarrow \cancel{T \leq}$$

$$\|\nabla f(x+\tau(y-x)) - \nabla f(x)\| \leq L\|\tau(y-x)\| \leq L\tau\|y-x\|$$

Lastly,

$$\boxed{|f(y)-f(x)-\langle \nabla f(x), y-x\rangle| \leq \int_0^1 L\tau\, d\tau \cdot \|y-x\|^2 = \frac{L}{2}\|y-x\|^2}$$

In order to prove that second defⁿ implies first, we have,

$$|f(y) - f(x) - \langle \nabla f(x), y-x \rangle| \le \frac{L}{2} \|x-y\|^2 \rightarrow (1)$$

~~f(y)~~ In the above inequality interchang x and y,

$$|f(x) - f(y) - \langle f(y), x-y \rangle| \le \frac{L}{2} \|x-y\|^2 \rightarrow (2)$$

Add (1) and (2)

$$|f(y) - f(x) - \langle \nabla f(x), y-x \rangle| + |f(x) - f(y) - \langle f(y), x-y \rangle|$$
$$\le L\|x-y\|^2$$

& using $|a+b| \le |a| + |b|$ y

$$|\langle \nabla f(x), \cancel{y-x}\ x-y \rangle - \langle \nabla f(y), x-y \rangle| \le L\|x-y\|^2$$

$$|(\nabla f(x) - \nabla f(y))^T (x-y)| \le L\|x-y\|^2$$

(7b) Given that f is a convex function, and $L = 2p$

To prove: $\frac{1}{4p} \|\nabla f(x) - \nabla f(x)\|^2 \leq |f(y) - f(x) - \nabla f(x)^T (y-x)| \leq p\|y-x\|^2$

Proof:- From the previous part of this question, we have

$$|f(y) - f(x) - \nabla f(x)^T (y-x)| \leq p\|y-x\|^2 \longrightarrow (1)$$

$\{\because L = 2p \Rightarrow \frac{L}{2} = p\}$

Now in order to prove left inequality;

~~Observe~~ $\because$ the function f is convex,

therefore $\forall\, x, y \in \cancel{} \mathbb{R}^n$

$$\boxed{f(y) - f(x) \geq f(x)^T (y-x)}$$

Therefore, above in (1) we don't need ~~mod~~ abs as it already ≥ 0.

we can write

$$f(x) - f(y) = \underbrace{f(x) - f(z)}_{\substack{\text{Using the} \\ \text{inequality for} \\ \text{convex functions}}} + \underbrace{f(z) - f(y)}_{\substack{\text{Using the} \\ \text{inequality} \\ \text{for L-smooth function.}}}$$

we have,

$$\begin{aligned}
f(x) - f(y) &\leq \nabla f(x)^T (x-z) + \nabla f(y)^T (z-y) + \frac{\cancel{2}p}{\cancel{2}} \|z-y\|^2 \\
&= \nabla f(x)^T (x-y+y-z) + \nabla f(y)^T (z-y) + p\|z-y\|^2 \\
&= \nabla f(x)^T (x-y) + (\nabla f(x)^T - \nabla f(y)^T) (y-z) + p\|z-y\|^2 \\
&= \nabla f(x)^T (x-y) + (\nabla f(x) - \nabla f(y))^T (y-z) + p\|z-y\|^2
\end{aligned}$$

Therefore, finally we have,

$$f(x) - f(y) \leq \nabla f(x)^T (x-y) + (\nabla f(x) - \nabla f(y))^T (y-z) + p\|z-y\|^2$$

Since ~~*~~ the above inequality is true for any x, y, z,

we ~~have~~, can put ~~$f(x)$~~ $\boxed{y - z = \frac{1}{2\rho}\left(\nabla f(x) - \nabla f(y)\right)}$

$$\therefore \quad f(x) - f(y) \le \nabla f(x)^T (x-y) + (\nabla f(x) - \nabla f(y))^T \times \frac{1}{2\rho}(\nabla f(x) - \nabla f(y)) + \frac{\cancel{\rho}\, \|\nabla f(x) - \nabla f(y)\|^2}{4\rho^{\cancel{2}}}$$

$$f(x) - f(y) \le \nabla f(x)^T (x-y) - \frac{1}{4\rho}\|\nabla f(x) - \nabla f(y)\|^2$$

$$\Rightarrow \quad \boxed{\frac{1}{4\rho}\|\nabla f(x) - \nabla f(y)\|^2 \le f(y) - f(x) - \nabla f(x)^T (y-x)} \rightarrow (2)$$

Using results (1) and (2)

we can write

$$\boxed{\frac{1}{4\rho}\|\nabla f(x) - \nabla f(y)\|^2 \le |f(y) - f(x) - \nabla f(x)^T (y-x)| \le \rho \|y-x\|^2}$$

Clarification :→ For the result (1), since $\nabla f(x)$ is Lipschitz-continuous we can write

$$\|\nabla f(x) - \nabla f(y)\| \le 2\rho \|x-y\|$$

and therefore the function f is also L-smooth.

Therefore, the function f is both L-smooth and convex