EE5606

Convex Optimization

Final Exam

K. Surya Prakash

Roll: EE18BTECH11026

Q1)

a) True.

Reason: A convex opt problem requires the minimising objective to be a convex func.

⇒ Need to prove:

local minima of a convex func ⇒ Global minima of the func.

$f: \mathbb{R}^d \to \mathbb{R}$. A convex func.

let $x^* \in D$ {domain of f} ⇒ Convex.

↳ local minima.

→ Let $y \in D \Rightarrow y - x^*$ is a feasible direction

Consider a small $\alpha > 0$ s.t

$f(x^*) \leq f(x^* + \alpha(y - x^*))$ } defines x^* local minima

$\Rightarrow f(x^*) \leq f(\cancel{x^*}\ \alpha y + (1-\alpha)x^*)$

$f(x^*) \leq \alpha f(y) + (1-\alpha) f(x^*)$ } ∵ func is convex.

$\alpha f(x^*) \leq \alpha f(y)$

$\boxed{f(x^*) \leq f(y)}$ $\Rightarrow \alpha \geq 0$

∴ for any arbitrary point y in D

$\Rightarrow f(x^*) \leq f(y)$

↳ Global minima

(b) True

$Epi(f) = \{(x,t) \; f(x) \le t\}$ $x \in Dom f$

Let f be a convex fun with $x, y \in Dom f$; $\alpha \in (0,1)$

$$\Rightarrow f(\alpha x + (1-\alpha) y) \le \underbrace{\alpha f(x) + (1-\alpha) f(y)}_{\text{From defn of convex fun}}$$

$$\le \alpha(t) + (1-\alpha) t$$

$$\boxed{f(\alpha x + (1-\alpha) y) \le t}$$

$\therefore (\alpha x + (1-\alpha) y, t) \in Epi(f)$.
$\in Dom f$: convex set

$\therefore$ $Epi(f)$ is a convex set

(c) True

From the defn of strict convex func

$$f(\alpha x + (1-\alpha) y) < \alpha f(x) + (1-\alpha) f(y)$$

$x, y \in Dom f$, $\alpha \in (0,1)$

Also can be written as without loss of generality

$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y)$

↳ loose inequality

$\underbrace{\hspace{6cm}}_{\text{Convex.}}$

---

d) False

Eg: Consider a affine func : $\nabla^2 f = 0$

Eg $f(x) = 5x$

$f(\alpha x + (1-\alpha) y) = \alpha f(x) + (1-\alpha) f(y)$

$\underbrace{\hspace{6cm}}_{\text{Both convex \& concave}}$

(e) True : similar to 1(a)

Let f be the quasi convex objective

Let $y \in D \Rightarrow y - x^*$: a feasible direction

$\Rightarrow$ Consider a small $\alpha > 0$, st ; & x^* is local optimum

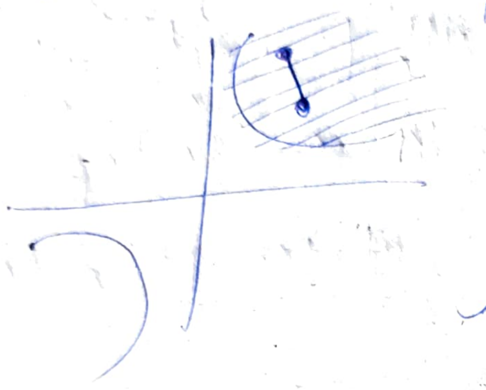
$$
\begin{aligned}
f(x^*) &\le f(x^* + \alpha(y - x^*)) \\
&\le f(\alpha y + (1-\alpha) x^*) \\
&\le \max(f(y), f(x^*)) \quad \text{: defn of quasi convexity} \\
&\le f(y) \quad (\because x^* \text{ is local opt})
\end{aligned}
$$

$\Rightarrow$ x^* is global optima for arbitrary 'y'

(f) True

[redacted]

$R = \{(x,y) \in \mathbb{R}^2 \mid xy \geq 1,\ x,y \geq 0\}$



The shaded regn is R

we can draw two points & all the pts on the connecting line segment lies in that region

---

(g) True

[redacted]

$$\max_{x}\ x_1 x_2 \cdots x_n \quad \text{st } \sum x_i = 1,\ x_i \geq 0 \quad \Rightarrow \quad \max_{x} \left((x_1 x_2 \cdots x_n)^{1/n} \right)^n \quad \text{st } \frac{\sum x_i}{n} = 1/n,\ x_i \geq 0$$

By the inequality $AM \geq GM$

$GM \leq AM.$

$$\Rightarrow \max_{x} \left(\frac{x_1 + \cdots x_n}{n} \right)^n = \left(\frac{1}{n}\right)^n =$$

$$\text{st } \frac{\sum x_i}{n} = 1/n \ ;\ x_i \geq 0$$

$\therefore$ obj. value $= \left(\frac{1}{n}\right)^n$

$\Rightarrow x_1 \cdot x_2 \cdots x_n = (1/n)^n$

&

$\sum x_i = 1$

can only happen if

$$\boxed{x_1 = x_2 = x_3 = \cdots x_n = \frac{1}{n}}$$

(h) False : Depends upon value of $f(x)$ & $f(y)$

* contradicting eg:

let $f(x)$: convex ; $g(y) = 0$ } convex

$h(x,y) = (f(x) + g(y))^2$

$$\frac{\partial^2 h(x,y)}{\partial x^2} = 2\left(f''(x) f(x) + (f'(x))^2\right) + 2 f''(x)\, g(y)$$

$$\frac{\partial^2 h(x,y)}{\partial y^2} = 2\left(g''(y) g(y) + (g'(y))^2 \right) + 2 g''(y)\, \cancel{g} f(x)$$

Considers

$h(x,y) \Rightarrow f(x) \le 0, \; g(y) \le 0$

$\therefore$ Then the trace of the Hessian can be negative $\Rightarrow$ Negative eigen values

$\Rightarrow$ Need not be convex.

Q2)

a) All are convex

All norm fun are convex

a) (2). $\min_{x} \|Ax - y\|_2^2 + \alpha \|x\|_2^2$

affine combination

$\rightarrow \|\cdot\|_2^2 \rightarrow$ convex func.

& No constraints on x

$\Rightarrow$ Convex op

(3) $\|x\|_2$ } convex func.

$\|A^T(y - Ax)\|_2^2 \le \alpha$

we know that

$A^T(y-Ax)$: Affine combinations

$\| \cdot \|_2^2 \rightarrow$ convex func

$\| A^T(y-Ax) \|_2^2 \leq \alpha \}$ Sublevel sets
$\Rightarrow$ convex

$\therefore$ obj + constraint both convex

4) $\min_x \|Ax-y\|_2^2 + \alpha\|x\|_1$
combination of convex functions, $\alpha \geq 0$
$\Rightarrow$ convex

(5) $\min_x \|x\|_1 \}$ convex objective

$\|A^T(y-Ax)\|_\infty \leq \alpha \}$ Sublevel set of a convex fun
$\Rightarrow$ convex

$\Rightarrow$ Hence a convex op

(b)

obj: $\|Ax-y\|_2^2 + \alpha\|x\|_2^2$

$= (Ax-y)^T(Ax-y) + \alpha x^T x$

$O(x) = (Ax)^T Ax + y^T y - y^T Ax - (Ax)^T y + \alpha x^T x$

$= x^T(A^TA)x + y^T y - y^T Ax - (Ax)^T y + \alpha x^T x$

$$\boxed{\frac{\partial O}{\partial x} = 2(A^TA)\underline{x} - 2A^T\underline{y} + 2\alpha \underline{x}.}$$

(c) Find x_1^* : optimum for (2).

* To find the optimum ⇒ min $O(x)$

⇒ $\frac{\partial O(x)}{\partial x} = 0$ } since $O(x)$ is a convex fun $\frac{\partial O(x)}{\partial(x)}$ gives the global minima

⇒ $2(A^TA)\underline{x} - 2A^Ty + 2\alpha\underline{x} = 0$

$(A^TA + \alpha I)\underline{x} = A^Ty$

$$\boxed{x_1^* = (A^TA + \alpha I)^{-1} A^Ty}$$

(2d)

Assuming α: large enough

Given $\|(A^TA+\alpha I)^{-1}A^T y\|_2 \le 1 \rightarrow$ (i)

& In (2c) we found that

optimum for (2)

$$\boxed{x_1^* = (A^TA+\alpha I)^{-1}A^T y}$$

& Eqn (i) describes that in this scenario

x_1^* lies inside a unit norm ball

$\Rightarrow \|x_1^*\|_2 \le 1$

Consider (3)

$\min \|x\|_2$

st $\|A^T(y-Ax)\|_2 \le \alpha$

$\Rightarrow$ Since α is large enough;

Constraint holds ~~for~~ $\forall\ x$ ~~irrespective of~~

$$\therefore\ x_2^* = \arg\min_{x_2} \|x_2\|_2$$

$$\Rightarrow \boxed{x_2^* = 0} \rightarrow \text{Origin}$$

$\Rightarrow \quad \|x_2^*\|_2 = 0$

thus, the constraint will be

$$\|A^T(y - \underline{0})\| = \|A^T y\| \leq \alpha$$

* <u>Note that</u>: $\|x_1^*\|$ need not equal to '$\underline{0}$'

This happens due to the presence of $\|Ax - y\|_2^2$ term in the objective, which tries to push $\underline{x}$ away from $\underline{0}$ to reduce this term.

<u>Hence</u>: $0 \leq \|x_1^*\|_2 \leq 1$

$\|x_2^*\|_2 = 0$

$$\Rightarrow \boxed{\|x_1^*\| \geq \|x_2^*\|}$$

* Equality holds when $\underline{y} = \underline{0}$

(e) Reformulating (5) as a LP

(5) $\min \|x\|_1$

st $\|A^T(y - Ax)\|_\infty \le \alpha$

Let $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$

x can be written as ~~sum~~ diff of two non-negative vectors. $:=$ $\boxed{x = x_+ - x_-}$

$\|x\|_1 = \underline{1}^T x_+ + \underline{1}^T x_-$; $\underline{1}$ unit vector $\in \mathbb{R}^n$

Reformulation:

$$\min_{x_+, x_-} \underline{1}^T x_+ + \underline{1}^T x_-$$

$$x_+ \ge 0 \; ; \; x_- \ge 0$$

$$x_+^T x_- = 0$$

$$A^T(y - A(x_+ - x_-)) \le \alpha(\underline{1})$$

$$-A^T(y - A(x_+ - x_-)) \le \alpha(\underline{1})$$

Explanation of constraint in next page

Breaking the constraint (Explanation)

$$\| A^T(y - Ax) \|_\infty \leq \alpha$$

Each element in the $n \times 1$ vector shd be less than α.

$\Rightarrow \max_{\text{all components}} (A^T(y - Ax)) \leq \alpha$

$\Rightarrow |(A^T(y - Ax))| \leq \alpha 1$

$\Rightarrow \boxed{-\alpha 1 \leq A^T(y - Ax) \leq \alpha}$

(4) Reformulating 4 as QP

(4) $\min_x \| Ax - y \|_2^2 + \alpha \| x \|_1$

$= (Ax - y)^T (Ax - y) + \alpha \| x \|_1$

$= x^T (A^T A) x - y^T A x - (Ax)^T y + y^T y$

$+ \alpha \| x \|_1$

* Reformulation:

$$\min_{x_+, x_-} \left\{ \begin{aligned} & (x_+ - x_-)(A^T A)(x_+ - x_-) \\ & - y^T A(x_+ - x_-) \\ & - (A(x_+ - x_-))^T y + y^T y \\ & + \alpha(\mathbf{1}^T x_+ + \mathbf{1}^T x_-) \end{aligned} \right.$$

$$x_+ \geq 0 \; ; \; x_- \geq 0$$

Explanation:

we can observe the obj of the form

$x^T Q x + p^T x + c$ $\Rightarrow$ this is a QP

while the constraints are linear

Q4)

a) Interpretation

For any Form, note that

$f_1(0) = f_2(0) = f_3(0) = f(0)$

$f_1(N-1) = f_2(N-1) = f_3(N-1) = f(N-1)$

@ $f_1(x) = \min_{m,c} mx + c$

$f(i) \leq mi + c$

For every x, will be on a line defined by a slope 'm' & intercept 'c'; such that all the other data points (samples) lie beneath / on the line

Algo : Start with $a = 0$

$$b = \left\{ \underset{i \in \{a+1, \dots N-1\}}{\arg\max} \; m_i = \left(\frac{f(i) - f(a)}{i - a} \right) \right.$$

$$c = \frac{f(a) b - a f(b)}{b - a} \Big\} \geq f(0)$$

If we find such a b

then $\forall x \in [a, b]$

$f_1(x) = m_i x + c = \left(\frac{f(b) - f(a)}{b - a} \right) x + c$

* $f_2(x) = \max_{\underline{\alpha}} \sum \alpha_i f(i)$

st $\sum i\alpha_i = x$

$\alpha \geq 0, \; \underline{1}^T\alpha = 1$

* Here too
for $x = 0 \Rightarrow \alpha_0 = 1 \Rightarrow f_2(0) = f(0)$

$x = N-1 \cdot \alpha_{N-1} = 1 \Rightarrow f_2(N-1) = f(N-1)$

$\Rightarrow \underline{1}^T\alpha = \sum \alpha_i = 1$

* $\sum i\alpha_i = x$: Expressing 'x' interms of data samples & interpolating them in the function.

* $f_2(x) = \max \sum \alpha_i f_i$

↳ We want to allot α_i more to f_i which is large $\Rightarrow$ we need to express x interms of $\alpha_i \Rightarrow \arg\max_i f_i$

→ Indirectly we want to expresses x in terms of a, b ; st slope. b/w a, b is highest

$$f_3(x) = \min_{\alpha} \sum \alpha_i f_i$$

$$\text{st } \sum i \alpha_i = x$$

$$\alpha \geq 0, \ \mathbb{1}^T \alpha = 1$$

Similar arguments follow as of (f_2).

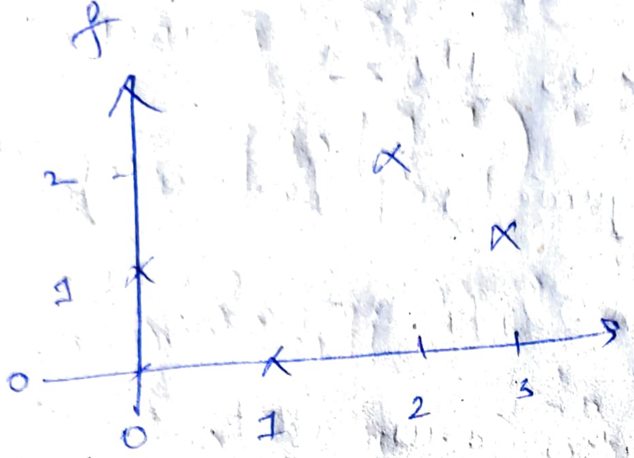
$f_3(x) = \min \sum \alpha_i f_i$

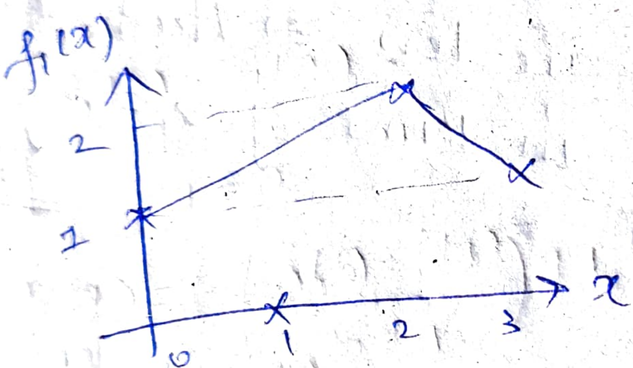
→ This makes a.w allot α_i more to f_i's which are less ⇒ Need to express x in terms of $\alpha_i = \arg\min_i f_i$

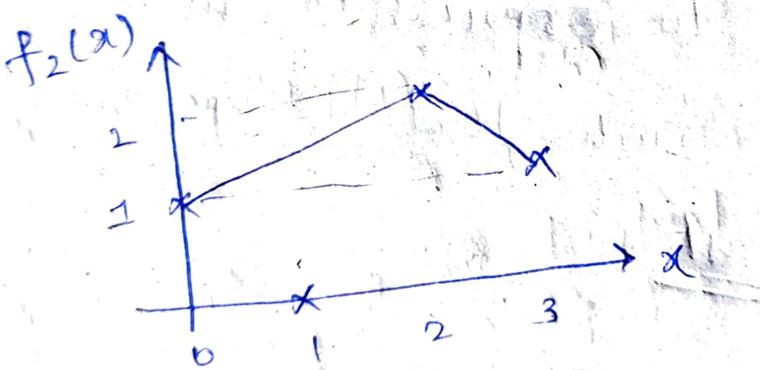
→ Want to express x in terms of (a, b) such that slope b/w a, b is least

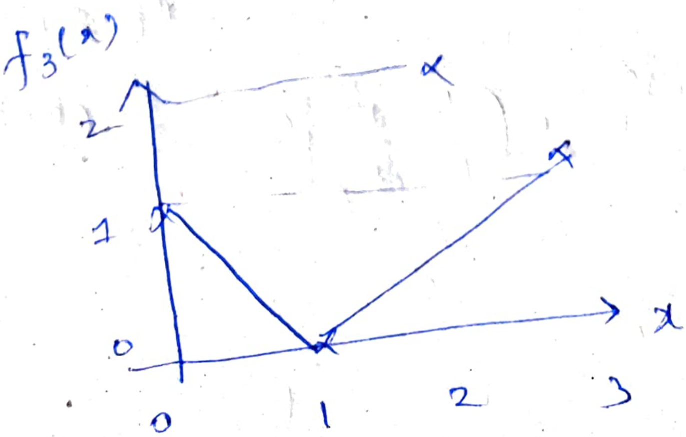
(b) Sketching f_1, f_2, f_3

[redacted]









c) f_1, f_2: ~~Convex~~ Concave

f_3: convex

Proof:

* In 4(d) we'll prove that $f_1(x) = f_2(x)$

* Equivalent representation of the same problem

⇒ Consider $f_2(x) = \max_{\alpha} \sum \alpha_i f_i$

$\sum \alpha_i = 1, \; \sum \alpha_i i = x; \; \alpha \geq 0$

where $x \in [0, N-1]$

* Let $x_1 = 0, \; x_2 = N-1$, pick a point

$\theta x_1 + (1-\theta) x_2 \in [0, N-1]$

* $f_2(\theta x_1 + (1-\theta) x_2) = \max_{\alpha} \sum \alpha_i f_i \rightarrow (1)$

$\sum \alpha_i i = x \quad , \sum \alpha_i = 1 \quad , \alpha \geq 0$

By simple observation one plausible α'

can be $\theta(0) + (1-\theta)(N-1) = x$

$$\Rightarrow \quad \alpha_i = \begin{cases} 0 & ; \; i \neq \{0, N-1\} \\ \theta & ; \; i = 0 \\ (1-\theta) & ; \; i = N-1 \end{cases}$$

for such config of α.

$$\sum \alpha_i f_i = \theta f(0) + (1-\theta) f(N-1))$$

we know that in $f_2(x)$

$$f_2(0) = f(0) \; ; \quad f_2(N-1) = f(N-1)$$

$$\sum \alpha_i f_i = \theta f_2(0) + (1-\theta) f_2(N-1) \rightarrow (2)$$

* we've shown that

$$f_2(\theta(0) + (1-\theta)(N-1)) = \max_{\alpha} \sum \alpha_i f_i \rightarrow (3)$$

ie, it checks all possible config of α & finds max of $\sum \alpha_i f_i$; where (2) is just one such configuration.

From (2) & (3)

$$\Rightarrow f_2(\theta(0) + (1-\theta)(N-1)) = \max_{\alpha} \sum \alpha_i f_i$$

$$\geq \sum \alpha_i f_i \quad \forall \text{ possible } \alpha$$

$$\geq \theta f_2(0) + (1-\theta) f_2(N-1)$$

$\Rightarrow f_2(\theta(0) + (1-\theta)(N-1)) \geq \theta f_2(0) + (1-\theta) f_2(N-1)$

Hence $\boxed{f_2 \text{ is CONCAVE}}$

* Since $f_1(x) = f_2(x)$ { proved in 4d }

$\Rightarrow$ $\boxed{f_1(x) \text{ is also concave}}$

---

* Consider $f_3(x) = \min \sum \alpha_i f_i$

$\sum \alpha_i = 1, \quad \sum i \alpha_i = x$

$\Rightarrow$ Consider $x_1 = 0, \; x_2 = N-1$

$f_3(\theta x_1 + (1-\theta) x_2) = \min \sum \alpha_i f_i$

* A plausible choise of α to satisfy the constraints can be

$$\alpha_i = \begin{cases} 0 & ; i \neq \{0, N-1\} \\ \theta & ; i = 0 \\ (1-\theta) & ; i = N-1 \end{cases}$$

$\Rightarrow$ For such α_i $\Rightarrow$ $\sum \alpha_i f_i = \theta f(0) + (1-\theta) f(N-1)$

* For $f_3(x)$: $f_3(0) = f(0)$; $f_3(N-1) = f(N-1)$

$\sum \alpha_i f_i = \theta f_3(0) + (1-\theta) f_3(N-1)$

* We know that as per the defn.

$$f_3(\theta|0) + (1-\theta)(N-1)) = \min \sum \alpha_i f_i$$

$$\leq \sum \alpha_i f_i \quad ; \forall \text{ possible } \alpha$$

$$\leq \theta f(0) + (1-\theta) f(N-1)$$

↳ One such 'α'

$$\Rightarrow f_3(\theta(0) + (1-\theta)N-1)) \leq \theta f(0) + (1-\theta) f(N-1)$$

$\Longrightarrow$ $\boxed{f_3 : \text{CONVEX}}$

(d) Prooving $f_1(x) = f_2(x)$

$$f_1(x) = \min mx + c$$
$$\text{st } f(i) \le mi + c \; ; \; \forall i \in \{0, \ldots N-1\}$$

$$f_2(x) = \max_{\alpha} \sum \alpha_i f(i)$$
$$\text{st } \sum i\alpha_i = x \; ; \quad \alpha \ge 0 ; \; 1^T\alpha = 1$$

* $f_1(x)$ can be reformulated as

$$a = 0$$

$$b = \arg\max_{i \in \{a+1, \ldots N-1\}} \frac{f(i) - f(a)}{i - a}$$

$$\Rightarrow \quad c = \frac{b f(a) - a f(b)}{b - a} \ge f(0)$$

$\Rightarrow$ for $x \in [a, b]$

$$\boxed{f_1(x) = \left(\frac{f(b) - f(a)}{b - a}\right) x + c}$$

Now $a = b$

& find another line (b)

Now assuming that $f_2(x) > f_1(x)$
then this formulation shd satisfy
all constraints

$\Rightarrow$

$$f_2(x) = \max_{\alpha} \sum \alpha_i f_i$$

$$\text{st} \quad \sum i \alpha_i = x \; ; \quad \alpha \geq 0 \, , \quad \sum \alpha_i = 1$$

* Let $f_2(x) = f_1(x)$

$a = 0$

$$\Rightarrow \quad f_2(x) = \left(\frac{f(b) - f(a)}{b-a} \right) x + c$$

$$\text{where} \quad c = \frac{b f(a) - a f(b)}{b-a}$$

$\{a, b\} \rightarrow$ make up the highest slope

Rearranging terms

$$\hookrightarrow \quad f_2(x) = f(b) \underbrace{\left(\frac{x-a}{b-a} \right)}_{\alpha_b} + f(a) \underbrace{\left(\frac{b-x}{b-a} \right)}_{\alpha_a}$$

$$\alpha_j = 0 \, , \quad i \neq \{a, b\}$$

$$= \alpha_a \, , \quad i = \cancel{\alpha} \, a$$

$$= \alpha_b \quad i = b$$

Need to check if α^i satisfy constraints

$$\Rightarrow \sum i\alpha_i = 0 + b\left(\frac{x-a}{b-a}\right) + a\left(\frac{b-x}{b-a}\right)$$

$$= x\left(\frac{b-a}{b-a}\right)$$

$$\boxed{\sum i\alpha_i = \underline{x}} \rightarrow \text{satisfied}$$

$$\Rightarrow 1^T\alpha = \sum \alpha_i = 0 + \left(\frac{x-a}{b-a}\right) + \left(\frac{b-x}{b-a}\right)$$

$$= \frac{b-a}{b-a}$$

$$\boxed{1^T\alpha = 1} \rightarrow \text{Satisfied}$$

$\because x \in [a, b]$, $x - a \geq 0$
$b - x \geq 0$

$\Rightarrow \alpha_a, \alpha_b \geq 0$

$\Rightarrow \boxed{\alpha_i \geq 0 \quad \forall i} \rightarrow$ Satisfied

Thus Our Assumption that $f_1(x) = f_2(x)$ is correct

e) Given m^*, c^* optimum of (7)

α^* " " (8)

& f_i contributes to $f_2(x)$

i.e, $f_2(x) = \sum_j \alpha_j^* f_j$

$\boxed{\alpha_i^* \neq 0}$

$\Rightarrow$ For f_i to contribute to $f_2(x)$

$\rightarrow$ i should be a sample with high slope, with which a line is drawn passing through x

$\Rightarrow \because \quad a \le x \le b$

$a, b \in \{0, 1, \ldots N-1\}$

either $a = i$ or $b = i$

$\rightarrow$ In such a case: $x \in [a, b]$

* $f_2(x)$: A line passes through a, b joining points

$$\Rightarrow \left.\begin{array}{l} f_2(a) = f(a) \\ f_2(b) = f(b) \end{array}\right\} \because i = \{a, b\}$$

$\Rightarrow \quad f_2(i) = f(i)$

$f_1(i) = f_2(i) = f(i)$

$\Rightarrow \boxed{f(i) = m^* i + c^*}$ //

Proving that if no 3 pts among $(i, f(i))$ are collinear $\Rightarrow$ d^* has atmost 2 non-zero entries.

* Let us assume that there exist 3 collinear points $(i, f(i))$

$(i_1, f(i_1))$; $(i_2, f(i_2))$; $(i_3, f(i_3))$

As proved earlier

Each $f(i_1) = m^* i_1 + c^*$

$f(i_2) = m^* i_2 + c^*$

$f(i_3) = m^* i_3 + c^*$

$\therefore$ For any $x \in [\min(i_1, i_2, i_3), \max(i_1, i_2, i_3)]$

i_1, i_2, i_3 can be contributors for x.

$\therefore$ $\alpha_{i_1} i_1 + \alpha_{i_2} i_2 + \alpha_{i_3} i_3 = x$

where, $\boxed{\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3} \neq 0}$ //

$\therefore$ If there exists 3 or more collinear Samples $\Rightarrow$ leads to more than 2 non-zero α_i^* entries.

$\Rightarrow$ If no 3 points are collinear

$\Rightarrow$ m^* for every pair of points

$x \in [a, b]$ $\Rightarrow$ has a unique m^*

where a, b contribute to x

$\Rightarrow$ Can have atmost 2 non-zero α entries

The End