Convex Optimization

ASSIGNMENT No. 02

EE21MTECH14002

Q.1

Sol: We have vector $\underline{y}$ and matrix A

Consider the following four optimization problems

$$\min_{\underline{x}} \|A\underline{x}-\underline{y}\|^2 + \alpha\|\underline{x}\|_2^2 \quad \text{—— (1)}$$

$$\min_{\underline{x}} \|A\underline{x}-\underline{y}\|_2^2 + \alpha\|x\|_1 \quad \text{—— (2)}$$

$$\min_{\underline{x}} \|\underline{x}\|_2 \quad \text{—— (3)}$$
$$\text{s.t } \|A^T(\underline{y}-A\underline{x})\|_2 \leq \alpha$$

$$\min \|\underline{x}\|_1 \quad \text{—— (4)}$$
$$\text{s.t } \|A^T(\underline{y}-A\underline{x})\|_\infty \leq \alpha$$

here $\alpha > 0$ is fixed

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

$$f(\underline{x}) = \|A\underline{x}-\underline{y}\|_2^2 + \alpha\|\underline{x}\|_2^2$$
$$= (A\underline{x}-\underline{y})^T(A\underline{x}-\underline{y}) + \alpha\underline{x}^T\underline{x}$$
$$= ((A\underline{x})^T-\underline{y}^T)(A\underline{x}-\underline{y}) + \alpha\underline{x}^T\underline{x}$$
$$= (\underline{x}^TA^T-\underline{y}^T)(A\underline{x}-\underline{y}) + \alpha\underline{x}^T\underline{x}$$
$$= \underline{x}^TA^TA\underline{x} - \underline{x}^TA^T\underline{y} - \underline{y}^TA\underline{x} + \underline{y}^T\underline{y} + \alpha\underline{x}^T\underline{x}$$
$$= \underline{x}^T(A^TA+\alpha I)\underline{x} - 2\underline{y}^TA\underline{x} + \underline{y}^T\underline{y}$$

To prove $f(\underline{x})$ is convex, we can prove its hessian is convex

$$f(\underline{x}) = 2A^TA\underline{x} + 2\alpha\underline{x} - 2A^T\underline{y}$$
$$= 2(A^TA+\alpha I)\underline{x} - 2A^T\underline{y}$$
$$\nabla^2 f = f''(x) = 2(A^TA+\alpha I)$$

$\because A^TA$ is PSD $\left\{ \begin{array}{l} \because x^T(A^TA)x = \|Ax\|_2^2 \geq 0 \\ \Rightarrow A^TA \text{ is PSD.} \end{array} \right.$

$\Rightarrow (A^TA + \alpha I)$ is PSD

$\Rightarrow f(x)$ is convex function

$\Rightarrow$ problem (1) is convex problem

(b)(ii) $\min_{x} \|Ax - y\|_2^2 + \alpha\|x\|_1$

$$f(x) = \underbrace{\|Ax - y\|_2^2}_{q} + \underbrace{\alpha\|x\|_1}_{p}$$

$p = \|x\|_1$ is L_1 norm and it is convex

To prove :- for convexity

$$f(\lambda v + (1-\lambda)w) \leq \lambda f(v) + (1-\lambda) f(w), \quad \lambda \in [0,1]$$

and from triangular inequality :-

$$\|\lambda v + (1-\lambda) w\|_1 \leq \|\lambda v\|_1 + \|(1-\lambda)w\|_1$$
$$= \lambda\|v\|_1 + (1-\lambda)\|w\|_1$$

$\Rightarrow p$ is convex

and to prove $f(x)$ is convex, we need to prove q is also convex

$$q(x) = \|Ax - b\|_2^2$$
$$= (Ax - b)^T(Ax - b)$$
$$= x^TA^TAx - 2b^TAx + b^Tb$$

$$q'(x) = 2A^TAx - 2b^TA$$
$$q''(x) = 2A^TA$$

A^TA must be PSD

$\because x^T(A^TA)x \Rightarrow \|Ax\|_2^2 \geq 0$

$\Rightarrow A^TA$ is PSD

$\Rightarrow \therefore q(x)$ is convex.

$$f(x) = q(x) + p(x)$$

$\because q(x)$ and $p(x)$ are convex

$\Rightarrow f(x)$ is also convex

(a)(iii)

$$\min_{\underline{x}} \|\underline{x}\|_2^2$$
$$\text{s.t.} \quad \|A^T(\underline{y} - A\underline{x})\|_2^2 \leq \alpha$$

Let $f(\underline{x}) = \|\underline{x}\|_2 = \underline{x}^T\underline{x}$

$f'(x) = 2\underline{x}$

$f''(x) = 2$

$\nabla^2 f = f''(x) = 2 \geq 0$, (Hessian func.) which is +ve

$\Rightarrow f(x)$ is convex.

As $A^T(\underline{y} - A\underline{x})$ is affine func. and norm of affine is convex.

$\|A^T(\underline{y} - A\underline{x})\|_2 \leq \alpha$ is sublevel set, which is also convex.

(a)(iv)

$$\min \|x\|_1$$
$$\text{s.t.} \quad \|A^T(\underline{y} - A\underline{x})\|_\infty$$

Let $f(\underline{x}) = \|\underline{x}\|_1$

$f(\underline{x})$ is convex as it is L_1 norm.

To prove :-

$$f(\lambda v + (1-\lambda) w) \leq \lambda f(v) + (1-\lambda) f(w), \lambda \in [0, 1]$$

and from triangle inequality.

$$\|\lambda v + (1-\lambda) w\|_1 \leq \|\lambda v\|_1 + \|(1-\lambda) w\|_1$$
$$= \lambda \|v\|_1 + (1-\lambda) \|w\|_1$$

$\Rightarrow f(\underline{x}) = \|\underline{x}\|_1$ is convex

Now, $\|A^T(\underline{y} - A\underline{x})\|_\infty = \max\{|a_i^T(y_i - a_i^T x_i)|\}, i = 1 \ldots$

∞ norm of affine func is convex

and $\|A^T(\underline{y} - A\underline{x})\|_\infty \leq \infty$ is sublevel ~~set~~ of $\|A^T(\underline{y} - A\underline{x})\|_\infty$ which is also convex

Hence, all 4 func. are convex

(b) find the gradient of objective of problem ①

Objective $f(\underline{x}) = \|A\underline{x} - \underline{y}\|_2^2 + \alpha \|\underline{x}\|_2^2$

$$= (A\underline{x} - \underline{y})^T (A\underline{x} - \underline{y}) + \alpha \underline{x}^T \underline{x}$$

$$= ((A\underline{x})^T - \underline{y}^T)(A\underline{x} - \underline{y}) + \alpha \underline{x}^T \underline{x}$$

$$= \underline{x}^T (A^T A + \alpha I)\underline{x} - 2\underline{y}^T A \underline{x} + \underline{y}^T \underline{y}$$

$$f'(\underline{x}) = 2(A^T A + \alpha I)\underline{x} - 2A^T \underline{y}$$

$\therefore$ gradient of objective of problem ①

$$\nabla f(\underline{x}) = 2(A^T A + \alpha I)\underline{x} - 2A^T \underline{y}$$

(c) From just above problem, we can see, gradient of objective

$$\nabla f = f'(\underline{x}) = 2(A^T A + \alpha I)\underline{x} - 2A^T \underline{y}$$

finding double derivative problem / Hessian -

$$\nabla^2 f = f''(\underline{x}) = 2(A^T A + \alpha I)$$

Now since $A^T A$ is PSD $\left(\because \underline{x}^T A^T A \underline{x} = \|A\underline{x}\|_2^2 \geq 0\right)$

$= A^T A + \alpha I$ is also PSD $(\alpha \geq 0)$

$\Rightarrow \nabla f(x) = f'(x) = 0$, will give optimum point for $\underline{x}$

$$\nabla f(x) = 2(A^T A + \alpha I)\underline{x} - 2A^T \underline{y} = 0$$

$$\boxed{\underline{x} = (A^T A + \alpha I)^{-1} A^T \underline{y}}$$

if x_1^* is optimum value, then

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

(d) Assume that α is large enough: in particular assume that

$$\|(A^T A + \alpha I)^{-1} A \underline{y}\|_2 \leq 1$$

If x_3^* is an optimum point of problem ③, show that

$$\|x_3^*\| \geq \|x_2^*\|_2$$

$$\min_{\underline{x}} \|\underline{x}\|_2$$

$$\text{s.t} \quad \|A^T(\underline{y} - A\underline{x})\|_2 < \alpha$$

Consider Constraints :- $\|A^T(\underline{y} - A\underline{x})\|_2 \leq \alpha$

Let A be and underdetermined matrix

$A \in \mathbb{R}^{m\times n}$ with $m<n$ (full-row-rank)

Consider $\underline{y} = A\underline{x}$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \cdots & \bar{a}_n \end{bmatrix}_{m\times n} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$m<n$

no. of eqn $\leq$ no. of unknown

Now, the problem can be formulated as -

$$\min \|\underline{x}\|_2$$

$$\text{s.t } A\underline{x} = y \;;\; A \in \mathbb{R}^{m\times n},\ m<n,\ x\in\mathbb{R}^n$$

$$\left(\text{i.e } \|A^T(\underline{y} - A\underline{x})\|_2 \leq \alpha \quad \alpha>0\right)$$

Lagrange func. :-

$$\mathcal{L}(\underline{x}, \lambda) = \|\underline{x}\|_2 + \lambda^T(\underline{y} - A\underline{x})$$

$$= \underline{x}^T\underline{x} + (\underline{y} - A\underline{x})^T\lambda$$

from optimization theory

$$\nabla \mathcal{L}(\underline{x},\lambda) = 0$$

$$\Rightarrow 2\underline{x} - A^T\lambda = 0$$

$$\Rightarrow \boxed{\underline{x} = 1/_2 A^T\lambda} \quad -①$$

as $\underline{y} = A\underline{x}$

$$= A\tfrac{1}{2}A^T\lambda$$

$$\underline{\lambda} = 2(AA^T)^{-1}\underline{y}$$

$$\boxed{\underline{x} = A^T(AA^T)^{-1}\underline{y}_n}$$

$$\Rightarrow \boxed{x_2^* = A^T (AA^T)^{-1} y}$$ for underdetermined matrix A

As $Ax = y$

$A(x - x_2^*) = 0$ $\{x - x_2^*$ lies in null space of $A\}$

Consider $(x - x_2^*)^T x_2^*$

$= (x - x_2^*)^T A^T (AA^T)^{-1} y$

$= 0$ $\{\because A(x - x_2^*) = 0\}$

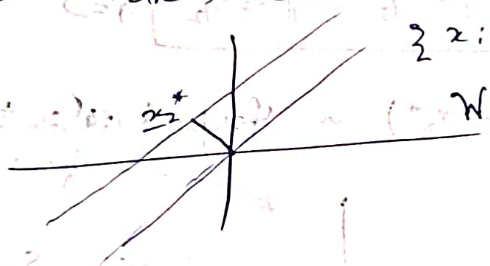
$$\Rightarrow \|x\|_2^2 = \|x_2^* + x - x_2^*\|_2^2$$

$$= \|x_2^*\|_2^2 + \|x - x_2^*\|_2^2 \geq \|x_2^*\|_2^2$$

$$\Rightarrow \|x_2^*\|^2 \leq \|x\|_2^2$$

$$\text{or } \boxed{\|x\|_2^2 \geq \|x_2^*\|_2^2} \quad -②$$

$\Rightarrow$ x_2^* has the smallest norm of any soln.



$\{x : Ax = y\}$, A is underdetermined $(m<n)$

$N(A) = \{x : Ax = 0\}$

x_2^* is perpendicular to Null space of A

$\Rightarrow$ $x_2^* \perp N(A)$:

Relation to problem ① :-

$$\min_x \|Ax - y\|_2^2 + \alpha \|x\|_2^2$$ and α is large enough

Now, if we assume $A \in \mathbb{R}^{m \times n}$, $m < n$

and $f_1(x) = \|Ax - y\|_2^2$, $f_2(x) = \|x\|_2^2$

we can find that from above interpretations –

$\|x\|_2^2$ get minimized when $y = Ax$

i.e. least norm soln minimizes

$$f_2(\underline{x}) = \|\underline{x}\|_2^2 \text{ with } f_1(\underline{x}) = 0$$

Minimizer of weighted sum objective -

$$f(\underline{x}) = f_1(\underline{x}) + \alpha f_2(\underline{x})$$
$$= \|A\underline{x} - \underline{y}\|_2^2 + \alpha\|\underline{x}\|_2^2$$

Soln is $\underline{x}_1^* = (A^TA + \alpha I)^{-1} A^T\underline{y}$

when $\alpha \to 0$

$\underline{x}_1^* \longrightarrow \underline{x}_2^*$

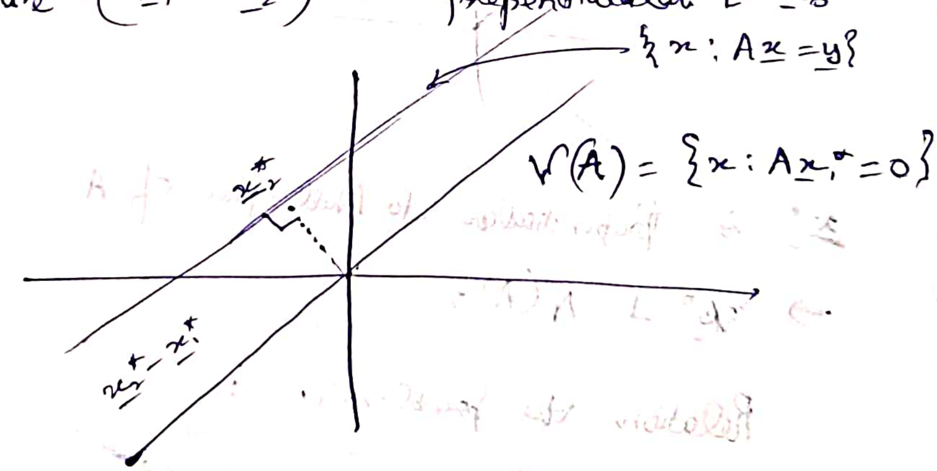
but $\alpha > 0$ and $\|(A^TA + \alpha I)^{-1} A^T\underline{y}\|_2 \leq 1$

$\therefore$ from eqn (2)

$$\|\underline{x}_1^*\|_2^2 \geq \|\underline{x}_2^*\|_2^2 .$$

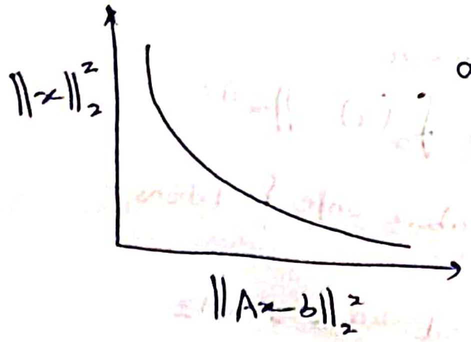
$$\Rightarrow \boxed{\|\underline{x}_1^*\|_2^2 \geq \|x_2^*\|_2} \text{ QED}$$

i.e $(\underline{x}_1^* - \underline{x}_2^*)$ is perpendicular to $\underline{x}_0^*$



$(\underline{x}_2^* - \underline{x}^*)$ perpendicular to $\underline{x}^*$

and $\boxed{\|\underline{x}_1^*\|_2 \geq \|\underline{x}_2^*\|_2}$



(e) Reformulate problem (4) as LP and problem 2 as QP :-

problem (4) as LP :-

$$\min_{x} \|x\|_1$$

$$\text{s.t. } \|A^T(y - Ax)\|_\infty \leq \alpha$$

$$\|x\|_1 = \underbrace{|x_1|}_{t_1} + \underbrace{|x_2|}_{t_2} + \cdots + \underbrace{|x_n|}_{t_n}$$

Equivalent LP is

$$\therefore \min. \; t_1 + t_2 + \cdots + t_n$$

$$\text{s.t } -t_1 \leq x_1 \leq t_2$$

$$\vdots$$

$$-t_n \leq x_n \leq t_n$$

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

$\because |x| \leq t$
$\Rightarrow -t \leq x \leq t$
and $\|x\|_\infty = \max\{|x_i|\}$
$\Rightarrow -\alpha \leq x_i \leq \alpha$
from $i = 1$ to n

So, LP is —

$$\min. \; 1^T x^+ + 1 x^-$$

$$\text{Sub to :- } -\alpha \leq A^T(y - A(x^+ - x^-)) \leq \alpha$$

$$x^+ \geq 0 \;, \; x^+ \geq 0$$

Variables : $x^+ \in \mathbb{R}^n, x^- \in \mathbb{R}^n$

problem 2 as QP :-

$$\min_{x} \|Ax - y\|_2^2 + \|x\|_1$$

$$\|x\|_1 = \underbrace{\|x_1\|}_{t_1} + \cdots + \underbrace{\|x_n\|}_{t_n} = \sum_{i=1}^{n} t_n$$

$$= 1^T t$$

$$\|Ax - y\|_2^2 = (Ax - y)^T (Ax - y)$$

$$= ((Ax)^T - y^T)(Ax - y)$$

$$= x^T(x^T A + \alpha I)x - 2y^T Ax + y^T y$$

Now,

$$\|Ax - y\|_2^2 + \|x\|_1$$

$$x^T(A^TA+\alpha I)x - 2y^TA\,x + y^Ty + 1^Tt \quad \text{—(a)}$$

Comparing (a) with Standard QP

$$f(x) = \frac{1}{2}x^TQx + \gamma^Tx + P$$

where Q is PSD

$$Q = 2(A^TA + \alpha I)$$

$$\gamma^T = -2y^TA$$

$$P = 1^Tt$$

$\therefore$ QP convex Problem for problem 2

$$\text{minimize } \frac{1}{2}x^TQx + \gamma^Tx + P$$

$$\underset{x}{\text{minimize}}\ x^T(A^TA + \alpha I)x - 2A^TAx + y^Ty + 1^Tt$$

$$\text{s.t.}\quad \left.\begin{array}{c} -t_1 \le x_1 \le t_1 \\ \vdots \\ -t_n \le x_n \le t_n \end{array}\right\} -t \le x \le t$$

Q.2

Sol:- (a) $f_1(x) = \sum_i \|A_ix - b_i\|_1$

Equivalent to the LP

$$\text{minimize } 1^Ts$$
$$\text{s.t.}\quad Ax - b \le s$$
$$Ax - b \ge -s$$

with variables $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^m$. Assume x is fixed in the problem, and we optimize only over s. The constraints say that

$$-s_k \le a_k^Tx - b_k \le s_k$$

for each k, i.e. $s_k \ge |a_k^Tx - b_k|$. The Objective func of the LP is separable so we achieve the optimum over s by choosing

$$s_k = |a_k^Tx - b_k|_3$$

and obtain the optimal value $f^*(x) = \|Ax-b\|$,

$\therefore$ optimizing over x and s simultaneously is equivalent to the original problem.

(b) $f_\infty(x) = \sum_i \|A_i x - b_i\|_\infty$

Equivalent to the LP

$$\begin{aligned} &\text{minimize} \quad t \\ &\text{s.t.} \quad Ax-b \leq t\mathbf{1} \\ &\qquad\quad Ax-b \geq -t\mathbf{1} \end{aligned}$$

in the variables $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, to see the equivalent, assume x is fixed in this problem and we ~~also~~ optimize only over t.

The constraints say that

$$-t \leq a_k^T x - b_k \leq t$$

for each k, i.e. $t \geq |a_k^T x - b_k|$

$$t \geq \max_k |a_k^T x - b_k| = \|Ax-b\|_\infty$$

Clearly, if x is fixed, the optimal value of the LP is $f^*(x) = \|Ax-b\|_\infty$. Therefore optimizing over t and x simultaneously is equivalent to the original problem.

(c) for $f_2(x)$:-

given $f_2(x) = \|Ax-b\|_2$

formulating it as SOCP

$$\min \quad t_1 + t_2 + t_3 + \dots + t_n$$

$$\text{st} \quad \|Ax_1 - b_1\|_2 < t_1, \ \|Ax_2 - b_2\|_2 < t_2 \dots$$

$\because$ Affine map preserves convexity, norm of affine map is convex if f is convex its ~~affine~~ alpha sublevel subsets are convex.

$$\therefore \|Ax_1 - b\|_2 < t \text{ is convex}$$

$$\therefore f_2(x) \text{ is convex}$$

Q.3

(a) To prove that the objective func. is quasi-concave we need to show that all α level superlevel sets are convex

$$\frac{\underline{\mu}^T \underline{x}}{\|V\underline{x}\|_2} \geq \alpha$$

$$\Rightarrow \|V\underline{x}\|_2 \leq \frac{1}{\alpha} \underline{\mu}^T \underline{x}$$

Given that V is symmetric which means quadratic part is convex and the above constraints look like SOCP constraints then it is convex.

(b)

$$\underline{z} = \frac{\underline{x}}{\underline{\mu}^T \underline{x}}$$

$$\Rightarrow \frac{\underline{z}}{\underline{1}^T \underline{z}} = \frac{\frac{\underline{x}}{\underline{\mu}^T \underline{x}}}{\frac{\underline{1}^T}{\underline{\mu}^T} \frac{\underline{x}}{\underline{x}}}$$

$$\Rightarrow \boxed{\underline{x} = \frac{\underline{z}}{\underline{1}^T \underline{z}}}$$

$$\frac{\underline{\mu}^T \underline{x}}{\|V\underline{x}\|_2} = \frac{\underline{\mu}^T \frac{\underline{z}}{\underline{1}^T \underline{z}}}{\left\|V \frac{\underline{z}}{\underline{1}^T \underline{z}}\right\|_2} = \frac{\operatorname{sgn}(\underline{1}^T \underline{z})}{\|V\underline{z}\|_2}$$

Given $\underline{\mu}^T \underline{x} \geq 0$ and $\underline{1}^T \underline{x} = 1$ which means $\underline{1}^T \underline{z} = \frac{1}{\underline{\mu}^T \underline{z}} \rightarrow 0$

$$\Rightarrow \frac{\underline{\mu}^T \underline{x}}{\|V\underline{x}\|_2} = \frac{1}{\|V\underline{z}\|_2}$$

$$\|\underline{x}\|_1 \leq L \Rightarrow \left\|\frac{\underline{z}}{\underline{1}^T \underline{z}}\right\|_1 \leq L$$

$$\|\underline{z}\|_1 \leq L \underline{1}^T \underline{z}$$

Now transformed problem is

$$\min \|V\underline{z}\|_2$$
$$\text{s.t } \|\underline{z}\|_1 \leq L\underline{1}^T\underline{z}$$
$$\underline{1}^T\underline{z} \geq 0$$

— The above transformed problem has both convex objective and constraints thus it is convex optimization problem.

Q.4
Soln (a) Lemma: $(A+B)^{-1} = A^{-1} - (I + A^{-1}B)^{-1} A^{-1} B A^{-1}$

If $g(x)$ is convex then so is $a^T g(x) a$ because the map linear with respective to $g(x)$. Now it is sufficient to prove the convexity of X^{-1}. We do this by contradiction assume that the function is not convex which means,

$$(\alpha A)^{-1} + ((1-\alpha)B)^{-1} < (\alpha A + (1-\alpha)B)^{-1}$$

$$\frac{1}{\alpha} A^{-1} + \frac{1}{1-\alpha} B^{-1} \leq \alpha A^{-1} - \frac{1-\alpha}{\alpha^2}\left(1 + \frac{1-\alpha}{\alpha} A^{-1}B\right)^{-1} A^{-1} B A^{-1}$$

Since the matrices are PSD multiplication on inequality will not change the sign

$$\frac{1}{1-\alpha} B^{-1} < \frac{1-\alpha}{\alpha^2}\left(1 + \frac{1-\alpha}{\alpha} A^{-1}B\right)^{-1} A^{-1} B A^{-1}$$

$$\Rightarrow \left(\frac{1-\alpha}{\alpha} A^{-1}B\right)^2 + \left(\frac{1-\alpha}{\alpha} A^{-1}B\right) + I < 0$$

Since $A \succeq 0$ and $B \succeq 0$ so is $A^{-1}B \succeq 0 \Rightarrow \frac{1-\alpha}{\alpha} A^{-1}B \succeq 0$

∴ Therefore the above obtained sum is just sum of PSD matrices which is PSD but we get negative definite which is a contradiction. Thus our assumption is wrong.

∴ X^{-1} is convex and so is $a^T X^{-1} a$

(b) Let $\overline{a}_i$ be ith column of identity matrix then from previous results $a_i^T X^{-1} a_i$ is convex, this function just picks (i,i) elements of X^{-1} which is a diagonal element. Thus diagonal elements are convex of combination of X.

(c) trace (X^{-1}) is sum of diagonal elements of X^{-1} which are individually convex. since sum of convex func. are convex, trace (X^{-1}) is convex

(d) Let transform this problem in epigraph problem.

$$\min_{t,X} \ t$$

$$\text{s.t.} \ t \geq \underline{a}^T X^{-1} \underline{a} \equiv \begin{bmatrix} t & \underline{a}^T \\ \underline{a} & X \end{bmatrix} \succeq 0$$

$$AX = B$$

$$X \succeq 0$$

Now we say $Z = \begin{bmatrix} X \\ \underline{v}^T \end{bmatrix}$, $\underline{v} = \begin{bmatrix} t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$$\begin{bmatrix} t & \underline{a}^T \\ \underline{a} & X \end{bmatrix} \succeq 0 \Rightarrow UZ + V \succeq 0$$

$$U = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} 0 & \underline{a}^T \\ \underline{a} & 0 \end{bmatrix}$$

$$AX = B \Rightarrow W.Z = B$$

$$W = [A \ 0]$$

$$X \succeq 0 \Rightarrow YZ \succeq 0$$

$$Y = [I \ 0]$$

final SDP Problem is,

$$\min_{t,X} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} Z$$

$$\text{s.t.} \begin{bmatrix} U \\ W \\ -W \\ Y \end{bmatrix} Z \geq - \begin{bmatrix} V \\ B \\ -B \\ 0 \end{bmatrix}$$

(e) The reformulation is similar to previous one we can rewrite trace as

$$\text{trace}(X^{-1}) = \sum_i \underline{a}_i^T X^{-1} \underline{a}_i$$

Where $\underline{a}_i$ is ith column of Identity matrix. Now we apply the same epigraph trick

$$\underline{a}_i^T X^{-1} \underline{a}_i \leq t_i$$

$$\begin{bmatrix} t_i & \underline{a}_i^T \\ \underline{a}_i & X \end{bmatrix} \succeq 0$$

$$\min_{t, X} \begin{bmatrix} \underline{0}^T & 1 \end{bmatrix} Z$$

$$\text{s.t.} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ \vdots \\ U_n \\ \tilde{W} \\ -\tilde{W} \\ Y \end{bmatrix} Z \geq - \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_n \\ B \\ -B \\ 0 \end{bmatrix}$$

where $Z = \begin{bmatrix} X \\ \underline{t}^T \end{bmatrix}$

Ar

Q.6

Consider the following program

5-4
6-3

$$\min_{x_1, x_2} \; x_1^2 + 2x_2^2 + x_1 x_2 - x_1$$

$$\text{s.t.} \quad x_1 - 2x_2 \leq u_1$$

$$x_1 + 4x_2 \leq -u_2$$

$$5x_1 + 76x_2 \leq 1$$

with variables x_1, x_2 and parameters u_1 and u_2

(a)

Standard QP

$$\min_{\underline{x}} \; \frac{1}{2} \underline{x}^T Q \underline{x} + \underline{x}^T \underline{x} + P$$

$$\text{s.t.} \quad A\underline{x} \leq b \quad \text{(inequality constraints)}$$

$$C\underline{x} \leq d \quad \text{(equality constraints)}$$

where Q is PSD for convex QP

Let $f(x_1, x_2) = x_1^2 + 2x_2^2 - x_1 x_2 - x_1$ — (1)

$$Q = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad \gamma = \begin{bmatrix} d \\ e \end{bmatrix}, \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(sym matrix)

then consider the —

$\frac{1}{2} \underline{x}^T Q \underline{x} + \underline{r}^T x + p = \frac{1}{2} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} d \\ e \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + p$

$= \frac{1}{2} a x_1^2 + b x_1 x_2 + \frac{1}{2} c x_2^2 + d x_1 + e x_2 + p$ — ②

from eqn ① and ②

$a = 2$, $b = -1$, $c = 4$, $d = -1$,

$$Q = \begin{bmatrix} a & b \\ b & d \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}$$

$$\underline{r} = \begin{bmatrix} d \\ e \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$|Q| = 7 > 0$

calculating eig values

$|Q - \lambda I| = 0$

$(2-\lambda)(4-\lambda) - 1 = 0$

$\Rightarrow 8 - 2\lambda - 4\lambda + \lambda^2 - 1 = 0$

$\lambda \geq 0$

$|Q| \geq 0$ and $\lambda \geq 0$

Q is PSD

$\Rightarrow \therefore$ given problem is QP convex

It is verified through cvxpy program by defining is_pos_check function.

(b) Using cvxpy, Solve this problem with parameter value $u_1 = -2$, $u_2 = -3$ to find the optimal primal variables x_1^*, x_2^* and optimal dual variables λ_1^*, λ_2^* and

Program is attached.

Minimized ~~/~~ optimal objective value = 7.44

Ans

(c) Verify KKT conditions holds for the optimal primal and dual optimal variables

It is verified through program

① Primal feasible

Optimal primal variables should satisfied the constraints

$$A\underline{x} \leq b$$

② Dual feasible –

$$\underline{\lambda}^* \geq 0$$

③ Complementary slackness :-

$$\sum \lambda_i^* f(\underline{x}^*) = 0$$

$$\Rightarrow \boxed{\underline{\lambda}^{*T}(A\underline{x}^* - b) = 0}$$

④ Derivative of Lagrangian = 0

Lagrange of QP can be written as :-

$$L(\underline{x}, \underline{\lambda}) = \frac{1}{2}\underline{x}^T Q \underline{x} + \underline{\lambda}^T (A\underline{x} - b) + \underline{\gamma}^T \underline{x} \quad -①$$

Defining Lagrangian dual fn :-

$$g(\lambda) = \inf_{\underline{x}} L(\underline{x}, \underline{\lambda})$$

finding derivative of ①

$$\nabla_{\underline{x}} L(\underline{x}, \underline{\lambda}) = Q\underline{x} + A^T \underline{\lambda} + \underline{\gamma}$$

Now, $\boxed{\nabla_{\underline{x}} L(\underline{x}^*, \underline{\lambda}^*) = Q\underline{x}^* + \underline{\gamma} + (A^T \underline{\lambda}^*) = 0}$

(d) Let $p^*(u_1, u_2)$ be the optimal value of the problem with parameters u_1, u_2, sketch some level curves for $p^*(u_1, u_2)$

Level curves are drawn in attached python program

(e) Yes, p^* is convex

$\because \; p^*(u_1, u_2) \geq p^*(-2,-3) - \lambda^{*T} b$

This can be proved as below -

From formed p^*

$$\min_{x} f(x) \quad \text{s.t. } Ax \leq b \quad \left\} \; f(x) = x_1^2 + 2x_2^2 - x_1x_2 - x_1, \quad u_1 = -2 \text{ and } u_2 = -3 \right.$$

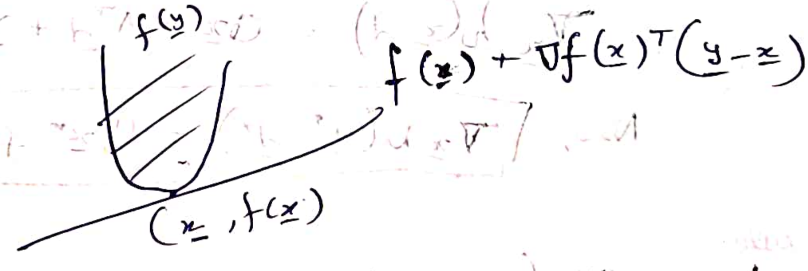
Problem with parameter u_1 and u_2

$$p^*(u_1, u_2) \begin{cases} \min_{x} f(x) \\ \text{s.t. } Ax \leq b' \end{cases}$$; here b' is with values u_1 and u_2

If $p^*(u_1, u_2)$ is optimal value of perturbed problem i.e. with parameter u_1 and u_2

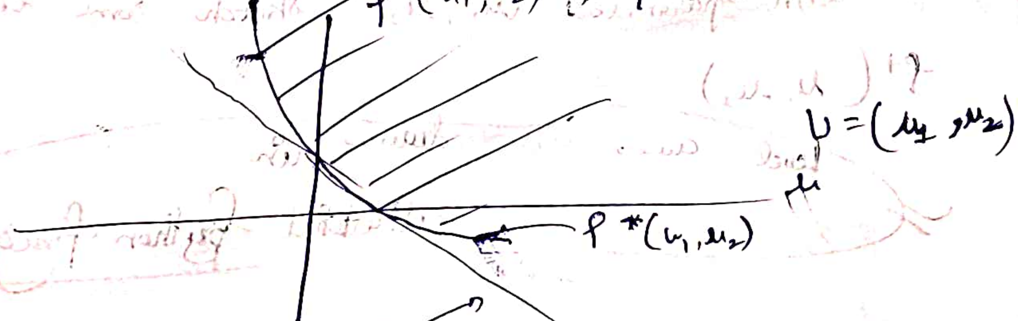
then strong duality holds (i.e. when original problem is convex and it satisfies subjected to condition

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$



Now, similarly drawing (A) -

$p^*(u_1, u_2) \geq p^*(-2,-3) - \lambda^{*T} b'$



So convexity of $p^*(u_1, u_2)$ can be varified from above epigraph curve.

Also, epigraph $= \{(u_1, u_2, t) \mid p^*(u_1, u_2) \leq t\}$

$= \{(u_1, u_2, t) \mid \exists\, x$ s.t.

$f(x) \leq t\ ,\ Ax \leq b'\}$

This is projection of convex set

$\{(u_1, u_2, t) \mid f(x) \leq t\ ,\ Ax \leq b'\}$

Hence, $p^*(u_1, u_2)$ is convex.

(6) (f) Compute the partial derivatives of p^* at $u_1 = -2$, $u_2 = -3$ and verify their relationship to the optimal dual variables λ_1^* & λ_2^*.

let $p^*(u_1, u_2)$ is differentiable at $u_1 = -2$, $u_2 = -3$.
Then, provided the strong duality holds, optimal dual variables are related to $p^*(u_1, u_2)$ as -

$$p^*(u_1, u_2) \geq p^*(-2,-3) - \lambda^{*T} b' \quad -(A)$$

$$y = mx$$

where m = slope $= -\lambda^*$
& intercept $= c = p^*(-2,-3)$

diff (A) w.r.to u_i -

$$\frac{\partial p^*(\underline{u})}{\partial u_i} \geq \frac{\partial p^*(-2,-3)}{\partial u_i} - \frac{\partial(\lambda^{*T} b')}{\partial u_i}$$

$$\frac{\partial p^*(\underline{u})}{\partial u_i} \geq 0 - \lambda_i^*$$

$$\boxed{\frac{\partial p^*(\underline{u})}{\partial u_i} = -\lambda_i^*}$$

Now, computing Numerically, partial derivatives -

1st w.r.t. u_1 - let $u_1 = -2+t$, $u_2 = -3$ ~~λ t=3/6/~~

$$\lim_{t \to 0} \frac{p^*(-2+t, -3) - p^*(-2,-3)}{t}$$

$$= \frac{\partial p^*(-2,-3)}{\partial u_1}$$

from eqn (A), inequality holds for $t \geq 0$ -

$$\frac{p^*(-2+t, -3) - p^*(-2,-3)}{t} \geq -\lambda_1^* \quad -(2)$$

for $t < 0$

$$\frac{p^*(-2+t, -3) - p^*(-2,-3)}{t} \leq -\lambda_1^* \quad -(3)$$

from (2) & (3) -

$$\frac{\partial p^*(-2,-3)}{\partial u_1} = -\lambda_1^*$$

$$\boxed{\frac{\partial p^*(-2,-3)}{\partial u_1} = -2.8649}$$

Now, finding w.r.t. u_2 -

let $u_1 = -2$ & $u_2 = -3+t$, $t \geq 0$

$$\lim_{t \to 0} \frac{p^*(-2,-3+t) - p^*(-2,-3)}{t} = \frac{\partial p^*(-2,-3)}{\partial u_2}$$

from eqn (A), Inequality holds for $t \geq 0$ -

$$\frac{p^*(-2,-3+t) - p^*(-2,-3)}{t} \geq -\lambda_2^* \quad -(4)$$

for $t \leq 0$ -

$$\frac{p^*(-2,-3+t) - p^*(-2,-3)}{t} \leq -\lambda_2^* \quad -(5)$$

from (4) & (5) -

$$\boxed{\frac{\partial p^*(-2,-3)}{\partial u_2} = -\lambda_2^*}$$

$$\Rightarrow \boxed{\frac{\partial p^*(-2,-3)}{\partial u_2} = -2.29803}$$

$\therefore$ Numerically & graphically, It is verified

$$\boxed{\frac{\partial p_i^*(u^*)}{\partial u_i} = -\lambda_i^*}$$

(5) You are given the samples of a function f at 0, 1, 2 --- N-1 ; you are given the values $f(0), f(1), f(N-1)$. Our goal is to interpolate to find the value of f on all points in the interval $[0, N-1]$, in such a way that the interpolation is convex.

we will assume that we are given an $x \in [0, N-1]$, and we try to construct the interpolated value $f(x)$. We come up with the following three potential techniques, resulting in 3 interpolated functions $f_1(x)$, $f_2(x)$ and $f_3(x)$, written as linear programs :-

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

$$\text{s.t. } f(i) \le mi + c \text{ for } i = 0, 1 \ldots N-1$$

$$f_2(x) = \max_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$

$$\text{s.t. } \sum_{i=0}^{N-1} i\alpha_i = x$$

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

$$f_3(x) = \min_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$

$$\text{s.t. } \sum_{i=0}^{N-1} i\alpha_i = 1,$$

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

(a) Given an Interpretation for the functions $f_1(x)$, $f_2(x)$ & $f_3(x)$.

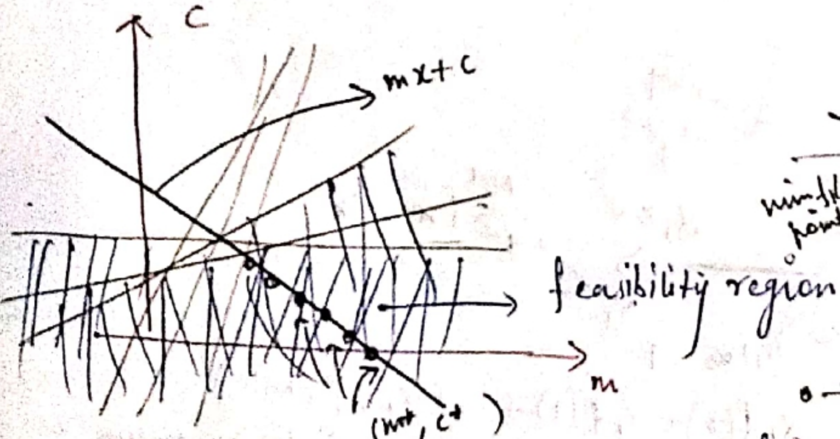
for $f_1(x)$ $\quad f_1(x) = \min\limits_{m,c} \; mx+c$

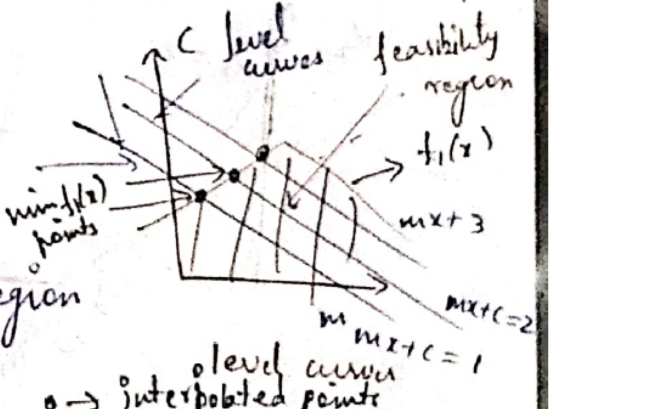
$\qquad\qquad s.t \cdot f(i) \le mi+c \quad , \text{ for } i = 0, 1 \ldots N-1$

drawing feasibility region :-

$$\left.\begin{aligned} f(0) &\le m\times 0 + c \\ f(0) &\le c \\ f(1) &\le m+c \\ f(2) &\le 2m+c \\ &\;\vdots \\ f(N-1) &\le (N-1)m + c \end{aligned}\right\}$$

$f(0), f(1) \ldots f(N-1)$ values are given
$(i, f(i))$, $i = 0, \ldots N-1$ are given





Now, the curve $mx+c$ is a line, intersection of $mx+c$ with feasibility region will give m & c values such that $\frac{f}{f_1}$ $mx+c$ is minimized, which give m^* & c^* for the problem $f_1(x)$

NOTE that : objective is linear function of m & c & constraints are also affine function of m & c.

$\Rightarrow$ $f_1(x)$ is convex optimization problem, as $f_1(x)$ objective is concave ~~&~~ in constraints are convex & it must be convex along the line $mx+c$.

for $f_2(x)$:-
$$f_2(x) = \max_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$
$$\text{s.t.} \sum_{i=0}^{N-1} i\alpha_i = x$$
$$\underline{\alpha} \geq 0, \ \underline{1}^T \underline{\alpha} = 1$$

Expanding the constraints, we get =

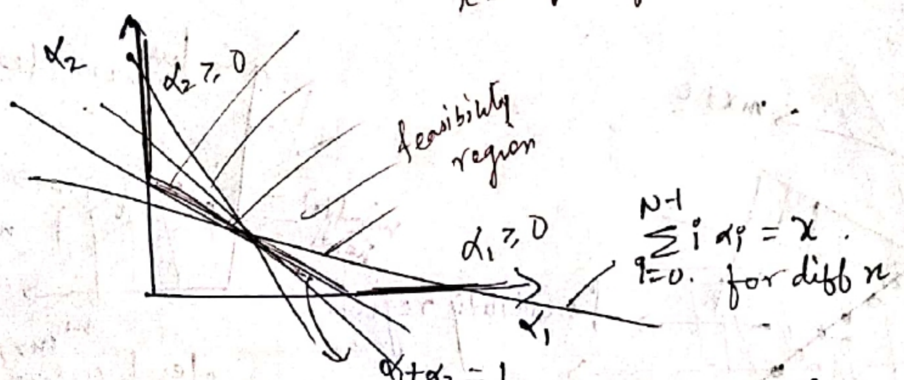
$\alpha_1 + 2\alpha_2 + \cdots + (N-1)\alpha_{N-1} = x$

$\alpha_0, \alpha_1, \alpha_2 \cdots \alpha_{N-1} \geq 0$

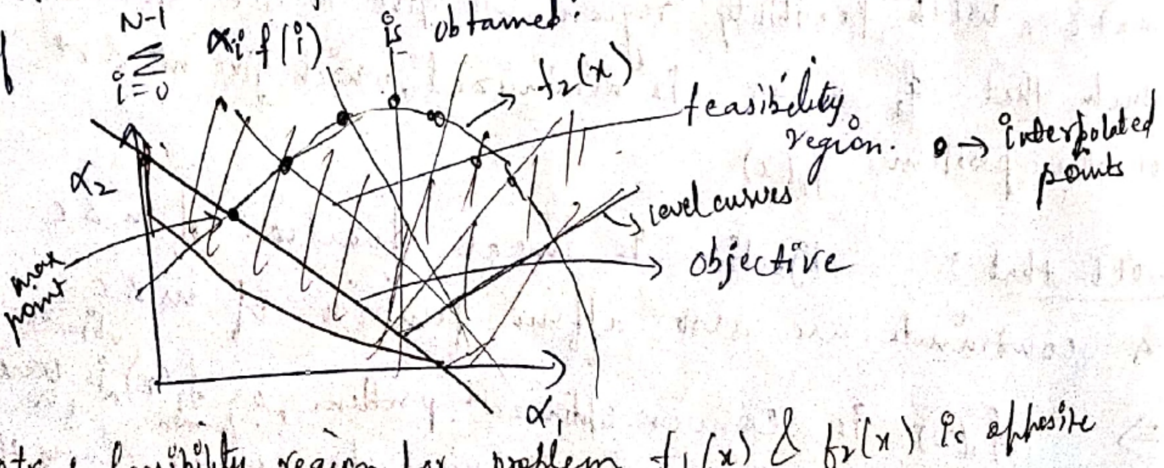
$\alpha_1 + \alpha_2 + \cdots \alpha_{N-1} = 1$

let's consider (α_1, α_2) only as $\underline{\alpha}$ for interpretation of $f_2(x)$

Now, drawing feasibility region i.e. plotting constraints, we get
x = fix for only $f_2(x)$



Now, objective $= \alpha_0 f(0) + \alpha_1 f(1) + \cdots + \alpha_{N-1} f(N-1)$ must intersects feasibility region such that, maximum of $\sum_{i=0}^{N-1} \alpha_i f(i)$ is obtained.



Note : feasibility region for problem $f_1(x)$ & $f_2(x)$ is opposite but max & min is found for $f_2(x)$ & $f_1(x)$ respective, so, convex curve, we get will be same. $f_2(x)$ objective is concave & it is convex opt. problem.

for $f_3(x)$ - $f_3(x) = \min_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$

s.t. $\sum_{i=0}^{N-1} i\alpha_i = x$

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

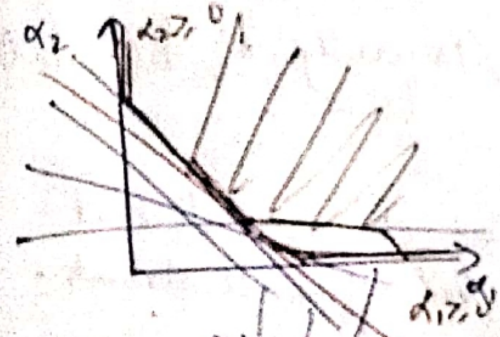
Expanding the constraints, we get -

$\alpha_1 + 2\alpha_2 + \dots + (N-1)\alpha_{N-1} = x, \ \alpha_0, \alpha_1, \dots \alpha_{N-1} \geq 0$

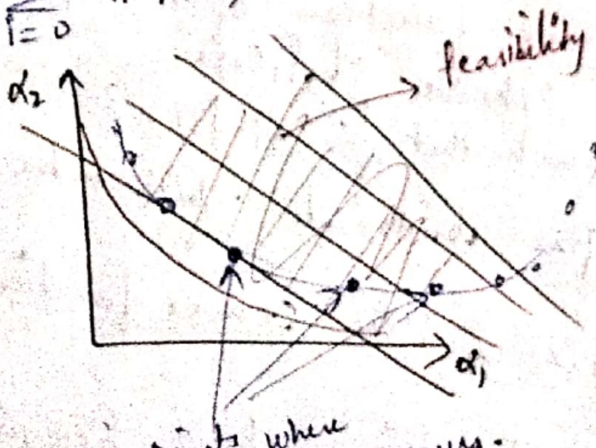
$\alpha_1 + \alpha_2 + \dots + \alpha_{N-1} = 1$

let's consider (α_1, α_2) only as $\underline{\alpha}$ for interpretation of $f_3(x)$

Now drawing feasibility region i.e plotting the constraints, we get -



Now, objective = $\alpha_0 f(0) + \alpha_1 f(1) + \dots + \alpha_{N-1} f(N-1)$, must intersects the feasibility region such that minimum of $\sum_{i=0}^{N-1} \alpha_i f(i)$ is obtained.

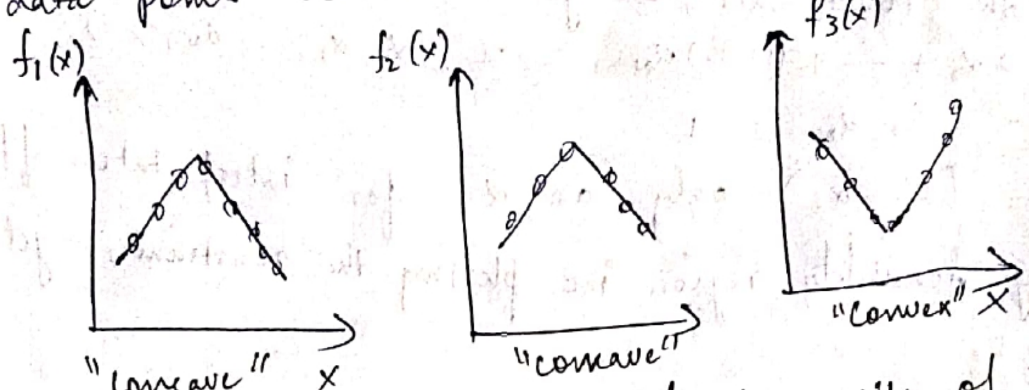


Joining the points i.e giving a convex function.

$\Rightarrow$ $f_3(x)$ is convex optimisation problem.

(b) for $N=4$, $f(0)=1, f(1)=0, f(2)=2, f(3)=1$,
sketch a plot of $f_1(x)$, $f_2(x)$ & $f_3(x)$.

Sketch of plot is made in python program for 100 data points between 0 to 3. (0 to N-1)



(c) Comment on the convexity / concavity of $f_1(x)$, $f_2(x)$ & $f_3(x)$ in general.

for $f_1(x)$ – m & c values are choosen such that – $mx+c$ is minimized intersecting the affine constraints.

$f_1(x)$ is concave function of x.

for $f_2(x)$ – α is choosen such that $\sum_{i=0}^{N-1} \alpha_i f(i)$ gets maximized intersecting the affine constraints which makes the feasibility region.

$f_2(x)$ is also a concave function of x. feasibility region of $f_2(x)$ is opposite of $f_1(x)$.

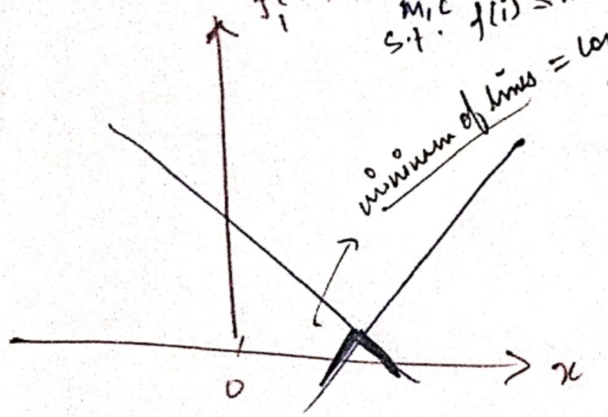
for $f_3(x)$ – α is choosen such that $\sum_{i=0}^{N-1} \alpha_i f(i)$ gets minimized intersecting the affine constraints which makes the feasibility region.

$f_3(x)$ is opposite of $f_2(x)$ & it is convex function of x.

let's go through each above functions $f_1(x)$, $f_2(x)$ & $f_3(x)$ through graphical interpretations, we get as –

e

$$f_1(x) = \min_{m,c} \; mx + c \quad \text{s.t. } f(i) \le mi + c$$

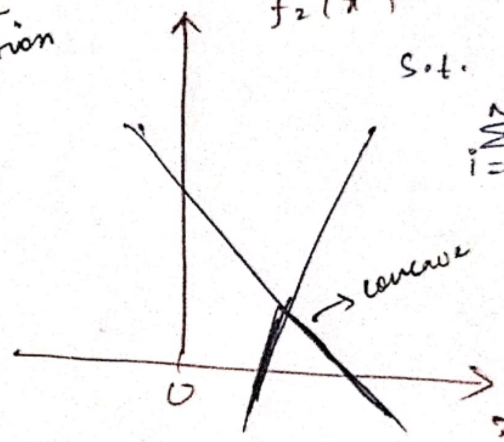


$$f_2(x) = \cancel{\min} \max_{\underline{\alpha}} \; \alpha_i f(i)$$

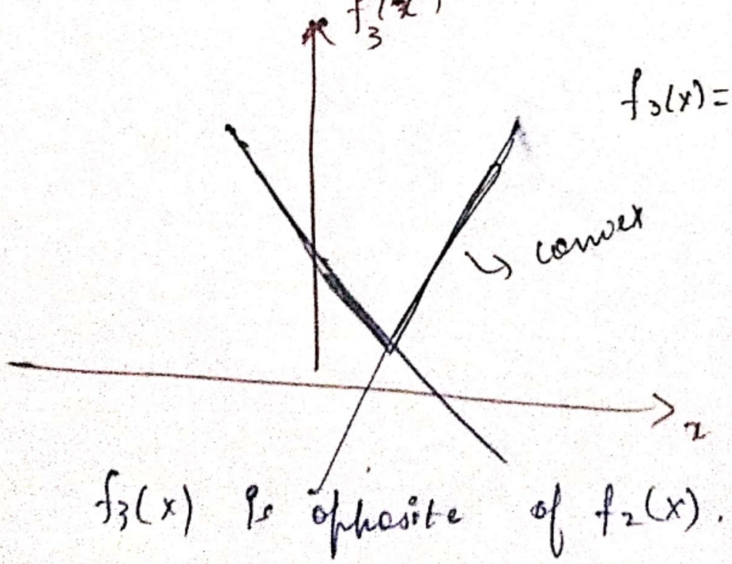
$$\text{s.t. } \mathbf{1}^T \underline{\alpha} = 1$$

$$\sum_{i=0}^{N-1} i \alpha_i = x$$

$$\underline{\alpha} \geq 0$$



$f_1(x) = f_2(x)$ (Dual form)



$$f_3(x) = \min_{\underline{\alpha}} \; \alpha_i f(i)$$

$$\text{s.t. } \mathbf{1}^T \underline{\alpha} = 1$$

$$\sum_{i=0}^{N-1} i \alpha_i = x$$

$$\underline{\alpha} \geq 0$$

$f_3(x)$ is opposite of $f_2(x)$.

for $f_1(x)$ $\quad f_1(x) = \min_{m,c} \; mx+c$

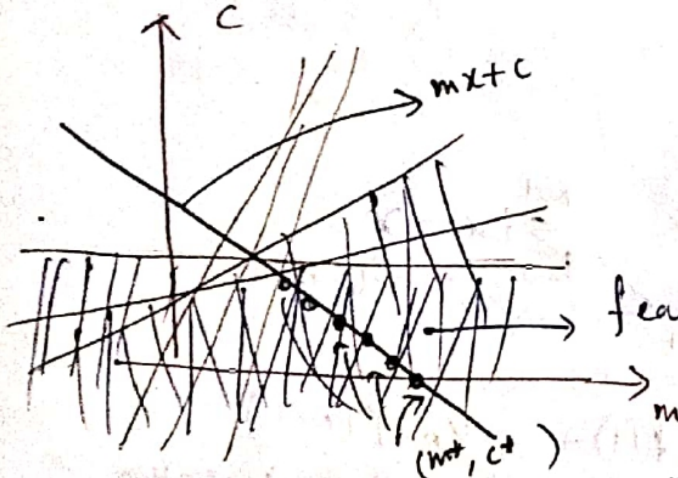
$\quad$ s.t. $f(i) \le mi+c$, for $i = 0,1 \ldots N-1$

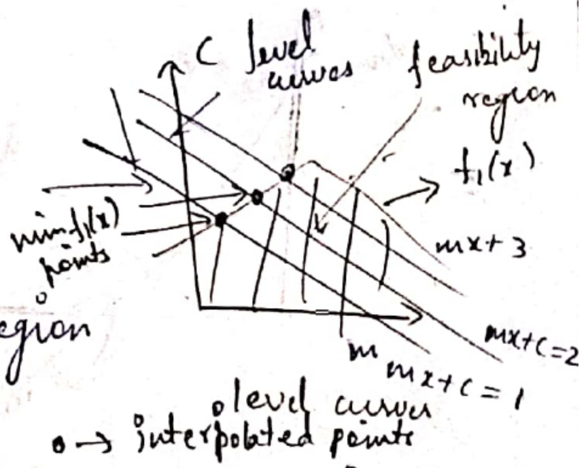
drawing feasibility region :-

$$
\left.\begin{aligned}
f(0) &\le m \times 0 + c \\
f(0) &\le c \\
f(1) &\le m + c \\
f(2) &\le 2m + c \\
&\vdots \\
f(N-1) &\le (N-1)m + c
\end{aligned}\right\}
$$

$f(0), f(1) \ldots f(N-1)$ values are given

$(i, f(i))$, $i = 0, \ldots N-1$ are given





Now, the curve $mx+c$ is a line, intersection of $mx+c$ with feasibility region will give m & c values such that ~~f_1~~ $mx+c$ is minimized, which give m^* & c^* for the problem $f_1(x)$

<u>NOTE that</u> : objective is linear function of m & c

& constraints are also affine function of m & c.

$\Rightarrow$ $f_1(x)$ is convex optimizatn problem, as ~~f_1~~ $f_1(x)$ objective is concave fn. & constraints are convex & it must be convex along the line $mx+c$.

for $f_2(x)$:-

$$f_2(x) = \max_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$

$$\text{s.t.} \quad \sum_{i=0}^{N-1} i\alpha_i = x$$

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

Expanding the constraints, we get :-

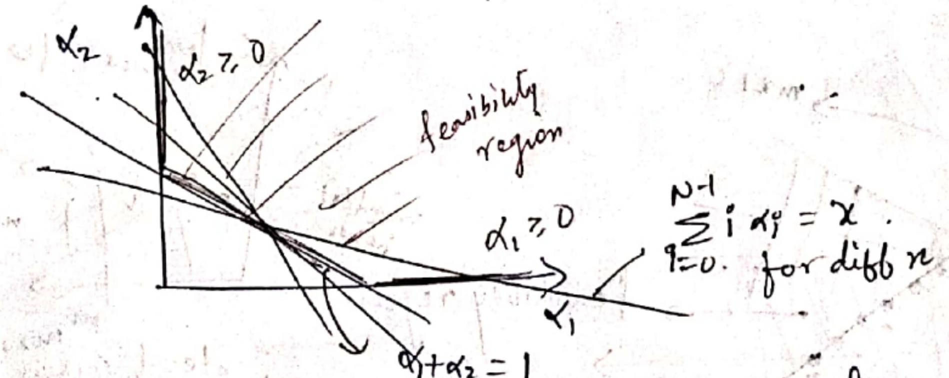
$\alpha_1 + 2\alpha_2 + \dots + (N-1)\alpha_{N-1} = x$

$\alpha_0, \alpha_1, \alpha_2 \dots \alpha_{N-1} \geq 0$

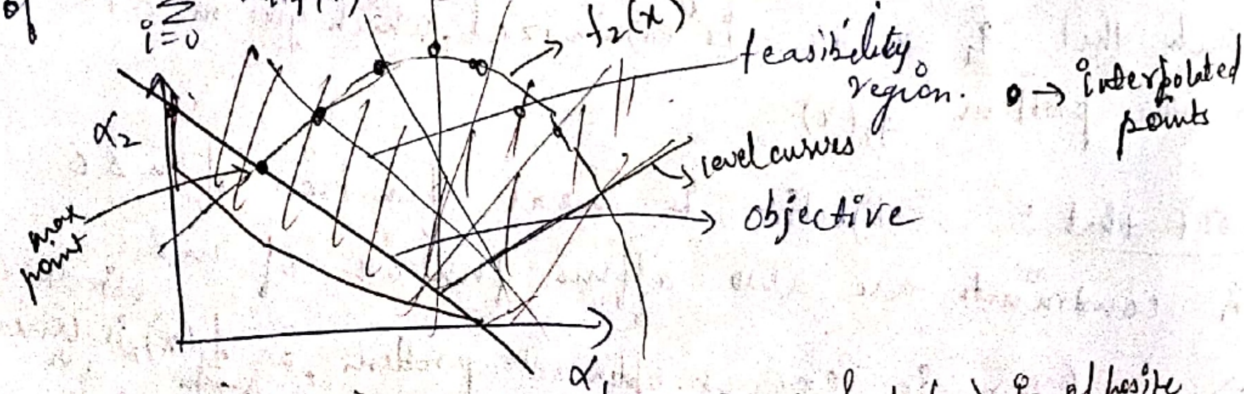
$\alpha_1 + \alpha_2 + \dots - \alpha_{N-1} = 1$

let's consider (α_1, α_2) only as $\underline{\alpha}$ for interpretation of $f_2(x)$

Now, drawing feasibility region i.e plotting constraints, we get
x = fix for only $f_2(x)$



Now, objective $= \alpha_0 f(0) + \alpha_1 f(1) + \dots + \alpha_{N-1} f(N-1)$ such that, maximum must intersects feasibility region is obtained.
of $\sum_{i=0}^{N-1} \alpha_i f(i)$



Note : feasibility region for problem $f_1(x)$ & $f_2(x)$ is opposite but max- & min is found for $f_2(x)$ & $f_1(x)$ respective, so, convex curve, we get will be same. $f_2(x)$ objective is concave & it is convex opt. problem.

for $f_3(x)$ - $$f_3(x) = \min_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$

$$\text{s.t.} \quad \sum_{i=0}^{N-1} i\alpha_i = x$$

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

Expanding the constraints, we get -

$\alpha_1 + 2\alpha_2 + \cdots + (N-1)\alpha_{N-1} = x$, $\alpha_0, \alpha_1, \ldots \alpha_{N-1} \geq 0$

$\alpha_1 + \alpha_2 + \cdots + \alpha_{N-1} = 1$

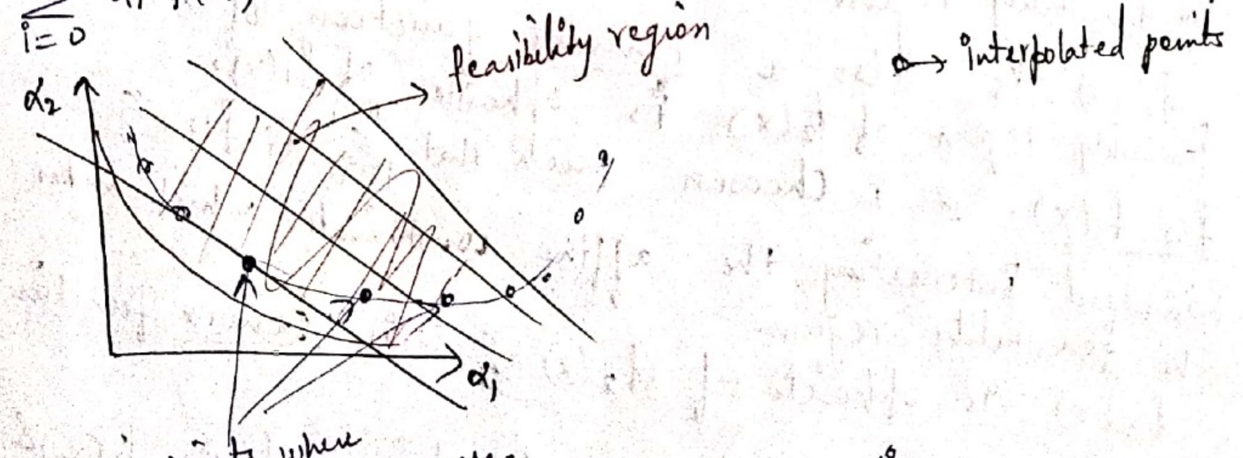
let's consider (α_1, α_2) only as $\underline{\alpha}$ for interpretation of $f_3(x)$

Now drawing feasibility region i.e plotting the constraints, we get -

$\sum_{i=0}^{N-1} i\alpha_i = x$, for diff. x

$\alpha_1 + \alpha_2 = 1$

Now, objective $= \alpha_0 f(0) + \alpha_1 f(1) + \cdots + \alpha_{N-1} f(N-1)$, must intersects the feasibility region such that minimum of $\sum_{i=0}^{N-1} \alpha_i f(i)$ is obtained.



Joining the points is giving a convex function.

$\Rightarrow$ $f_3(x)$ is convex optimisation problem.

Scanned with CamScanner

(d) prove that (in general) $f_1(x) = f_2(x)$.

$$f_1(x) = \min_{m,c} mx + c$$
$$\text{s.t. } f(i) \leq mi + c$$
$$\text{for } i = 0, 1, \ldots N-1$$

Lagrangian function of $f_1(x)$ –, with dual variable $\underline{\alpha} \geq 0$

$$L(m, c, \underline{\alpha}) = mx + c + \sum_{i=0}^{N-1} \alpha_i \left(f(i) - (mi + c)\right)$$

Dual problem can be written as – for dual variable $\alpha \geq 0$ –

$$D = \max_{\alpha} g(\underline{\alpha})$$
$$\text{subject to } \alpha \geq 0$$

where $g(\alpha) = \inf_{m,c} L(m, c, \underline{\alpha})$

finding $g(\alpha)$ –

$$g(\alpha) = \inf_{m,c} L(m, c, \underline{\alpha})$$
$$= \inf_{m,c} \left(mx + c + \sum_{i=0}^{N-1} \alpha_i \left(f(i) - (mi + c)\right)\right)$$

derivatives of $L(m, c, \alpha)$ w.r.t m & c must be equal to 0 – for dual variable $\underline{\alpha} \geq 0$ –

$$\frac{\partial L(m, c, \alpha)}{\partial m} = x - \sum_{i=0}^{N-1} \alpha_i \cdot i = 0$$

$$\Rightarrow \boxed{\sum_{i=0}^{N-1} \alpha_i \cdot i = x} \quad \text{— (1)}$$

Also, $$\frac{\partial L(m, c, \alpha)}{\partial c} = 1 - \sum_{i=0}^{N-1} \alpha_i = 0$$

$$\Rightarrow \boxed{\sum_{i=0}^{N-1} \alpha_i = 1} \quad \text{— (2)}$$

Rerewritting the dual D for $f_1(x)$ —

$$D = \max_{\underline{\alpha}} \inf_{m,c} L(m, c, \underline{\alpha})$$

subject to - $\underline{\alpha} \geq 0$

$$\sum_{i=0}^{N-1} i\,\alpha_i = x$$

$$\& \quad \sum_{i=0}^{N-1} \alpha_i = 1$$

$$\underline{\alpha} \geq 0 \quad \text{(Dual feasible)}$$

$$\cancel{\max_{\alpha}} \inf_{m,c} L(m, c, \alpha) = g(\alpha)$$

$$= \inf_{m,c} \left(mx + c + \sum_{i=0} \alpha_i \left(f(i) - (mi + c) \right) \right)$$

$$= \inf_{m,c} \Big(\sum_{i=0}^{N-1} \cancel{m} i\alpha_i - \sum_{i=0}^{N-1} \cancel{\alpha_i} m i + \sum_{i=0}^{N-1} \alpha_i f(i)$$
$$- c \sum_{i=0}^{N-1} \alpha_i + c \Big)$$

$$= \inf_{m,c} \left(\sum_{i=0}^{N-1} \alpha_i f(i) - c + c \right) \quad \left\{ \because \sum_{i=0}^{N-1} \alpha_i = 1 \ \& \ \sum_{i=0}^{N-1} i\,\alpha_i = x \right.$$

$$= \inf_{m,c} \left(\sum_{i=0}^{N-1} \alpha_i f(i) \right)$$

$$g(\alpha) = \sum_{i=0}^{N-1} \alpha_i f(i).$$

Now, dual of $f_1(x)$ can be written as —

$$D = \max_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$$

subj to - $\sum_{i=0}^{N-1} \alpha_i = 1$

$$\sum_{i=0}^{N-1} i\,\alpha_i = x, \quad \underline{\alpha} \geq 0$$

D is nothing but $f_2(x)$.

$\therefore \quad f_2(x) = D = \min_{\underline{\alpha}} \sum_{i=0}^{N-1} \alpha_i f(i)$

$$\text{subj to.} \quad \underline{1}^T \underline{\alpha} = 1$$

$$\sum_{i=0}^{N-1} i\alpha_i = x$$

$$\underline{\alpha} \geq 0$$

Slater's condition. If $f_1(x)$ is convex optimization problem satisfies Slater's condition if there exists m, c such that all constraints are satisfied strictly, then strong duality holds.

for $f_1(x)$:- constraints -

$$f(i) < mi + c \quad \text{for } i = 0, 1, \dots N-1$$

$\Rightarrow$ strong duality holds for $f_1(x)$..

$\Rightarrow$ optimum of $f_1(x)$ = optimum of D

$$\Rightarrow \quad f_1^*(x) = D^*$$

$$\Rightarrow \quad \boxed{f_1(x) = f_2(x)} \quad \text{as } D \text{ is } f_2(x) \text{ only.}$$

s strong duality holds, the function will satisfy all KKT conditions.

KKT conditions. suppose $\tilde{m}, \tilde{c}, \tilde{\alpha}$ satisfy the following

a) $\tilde{m}, \tilde{c}$ are primal feasible

$$f(i) \leq \tilde{m} i + \tilde{c} \quad , i = 0, 1 \dots N-1$$

b) $\tilde{\underline{\alpha}} \geq 0$, Dual feasible.

c) $\tilde{\alpha}(f(i) - \tilde{m} i + c) = 0$ (Complementary slackness)

d) gradient of Lagrangian $= 0$

$$\cancel{\nabla f_1(\tilde{m} i + c)}\ \nabla L(\tilde{m}, \tilde{c}, \tilde{\underline{\alpha}}) = 0$$

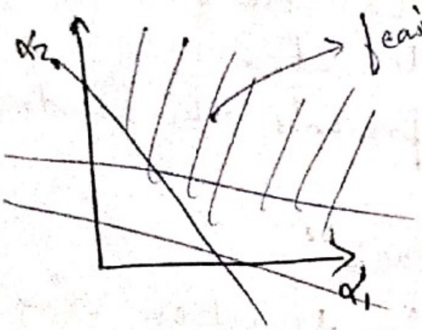
(e) Given an x, suppose m^*, c^* are optima of (7) & α^* is the optima of $f_2(x)$. If the value of $f(i)$ $d_i(x)$ that contributes to the value of $f_2(x)$, then show that $f(i) = m^* i + c^*$. Using this argue that if no three points among $(i, f(i))$ are collinear, then $\underline{\alpha}^*$ can have at most two non-zero entries.

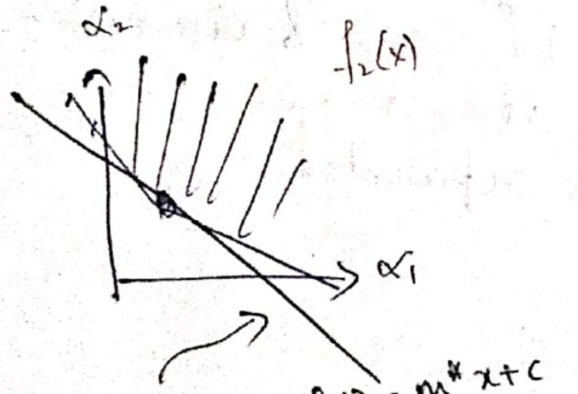
$m^*, c^* \longrightarrow$ optimum of $f_1(x)$

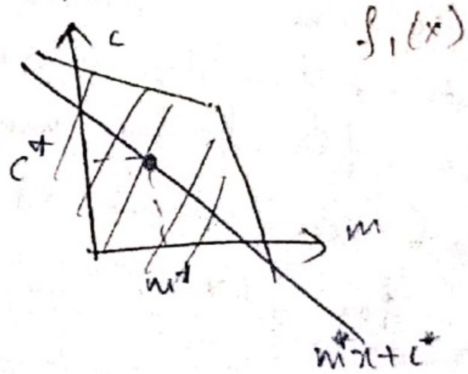
$\underline{\alpha}^* \longrightarrow$ optimum of $f_2(x)$

from part (d), we see that $f_2(x)$ is the dual form of $f_1(x)$.

$\Rightarrow$ $f(i) = m^* i + c^*$ for $f_2(x)$







As strong duality holds for $f_1(x)$, this implies all the KKT conditions will holds! m^*, c^*, α^* as primal & dual feasible variable

complementory slackness – for m^*, c^*

$$\sum_{i=0}^{N-1} \alpha_i^* (f(i) - m^* i + c) = 0$$

$$\Rightarrow \quad \alpha_i^* (f(i) - m^* i + c) = 0 \quad , \; i = 0, \text{to } N-1$$

As, we know,

(e) if $(f(i) - (mi+c)) \le 0$, then $\alpha_i = 0$ — ①

$i = 0$ to $N-1$

This interprets as if there is slackness in primal, then no slackness will be there for dual variable α_i^* & vice-versa.

or if $\alpha_i^* > 0$, then $f(i) - (m^* i + c^*) = 0$ — ②

As, we know from constraint of $f_2(x)$

$\alpha_i^* \geq 0$

$\Rightarrow$ ② will hold for complementary slackness KKT conditions.

$\Rightarrow$ $f(i) - (m^* i + c^*) = 0$

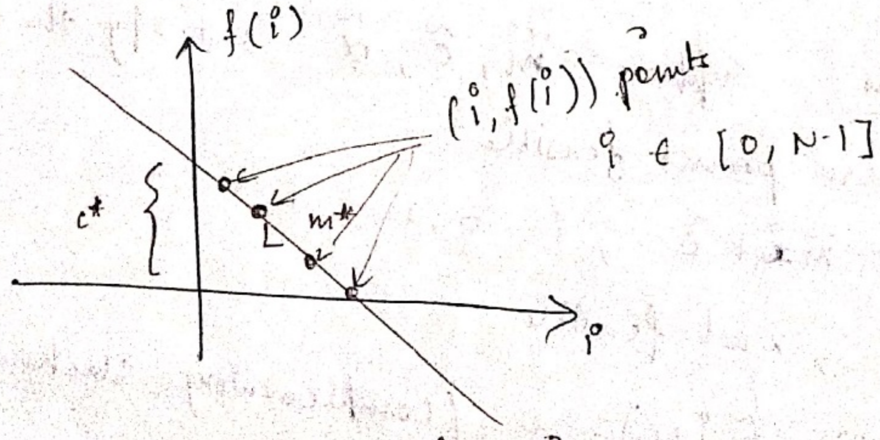
$\Rightarrow$ $\boxed{f(i) = m^* i + c^*}$

Hence proved.

This $f(i)$ contributes to value of $f_2(x)$ & it is equal to $\boxed{f(i) = m^* i + c^*}$.

Now, since $f(i) = m^* i + c^*$

$\Rightarrow$ all the points $(i, f(i))$ will lie on the line $f(i) = m^* i + c^*$



But this $f(i) = m^* i + c^*$ holds only when $\alpha_i^* > 0$ (Non-Zero Value) from ② - complementary slackness.

$\Rightarrow$ Number of non-zero α values is equal to number of collinear points $(i, f(i))$ that lies on $f(i) = mi + c$

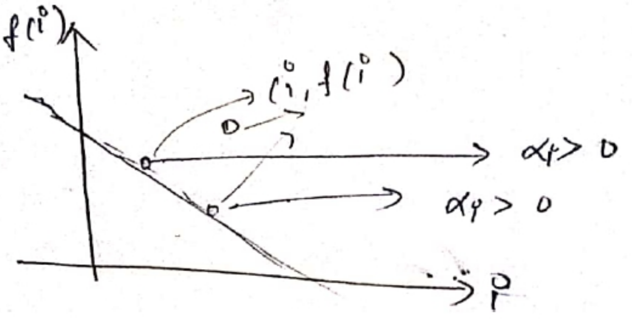
$\Rightarrow$ # non-zero α entries = # collinear $(i, f(i))$ points

<u>let us suppose</u> -

No 3 points $(i, f(i))$ are collinear

$\Rightarrow$ we can have atmost 2 collinear points $(i, f(i))$.

$\Rightarrow$ we can have atmost 2 non-zero α^* entries.



$\alpha_i \geq 0$ for atmost 2 times.