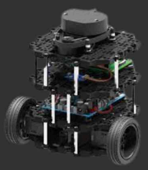
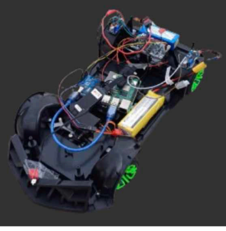
Safety Critical Control for Robotic Systems

By Manan Tayal

Lecture 2
Constrained Optimization


1.5x - 2.x

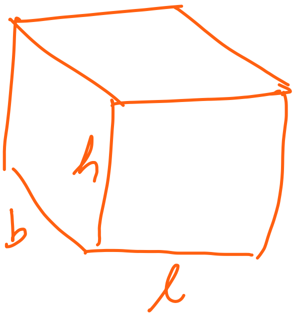
In this lecture:

- Optimization meaning & Notation
- Feasible & Descent Directions
- K K T Conditions (First-Order)
- Examples

Note: Optimization is very vast!

What is Optimization?

The act of making the best or most effective use of a situation or resource is Optimization.



$$\boxed{\begin{array}{l} \boxed{\min \; 2(lb + bh + hl)} \\ \boxed{\begin{array}{ll} \text{s.t} & lbh = V \\ & l > 0 \\ & b > 0 \\ & h > 0 \end{array}} \end{array}}$$

✓

Constrained Optimization Problem

$$\min \boxed{f(x)} \leftarrow \text{Objective } f^n$$

$$\boxed{\begin{array}{lll} \text{s.t.} & h_j(x) \le 0 & j = 1, 2, \dots\, l \\ & e_i(x) = 0 & i = 1, 2 \dots\, m \\ & x \in S & \end{array}} \leftarrow \text{Constraint set / Feasible set}$$

$h_j : \mathbb{R}^n \to \mathbb{R} \leftarrow$ Inequality Constraint fn

$e_i : \mathbb{R}^n \to \mathbb{R} \leftarrow$ Equality Constraint fn.

$f(x), h_j, e_i \leftarrow$ Sufficiently Smooth.

Feasible Set : $\left\{ x \;\middle|\; x \in S,\; h_j(x) \le 0,\; e_i(x) = 0 \;\; \forall i \in \{1, \dots m\} \;\&\; j \in \{1, \dots l\} \right\} \triangleq X$

$\min f(x)$ s.t $x \in X$

Global & Local Minima

A point $x^* \in X$ is said to be a global minima of $f(x)$ over X if $\boxed{f(x) \geq f(x^*)}$ for $x \in X$.

$f(x) > f(x^*) \rightarrow$ strict global minima

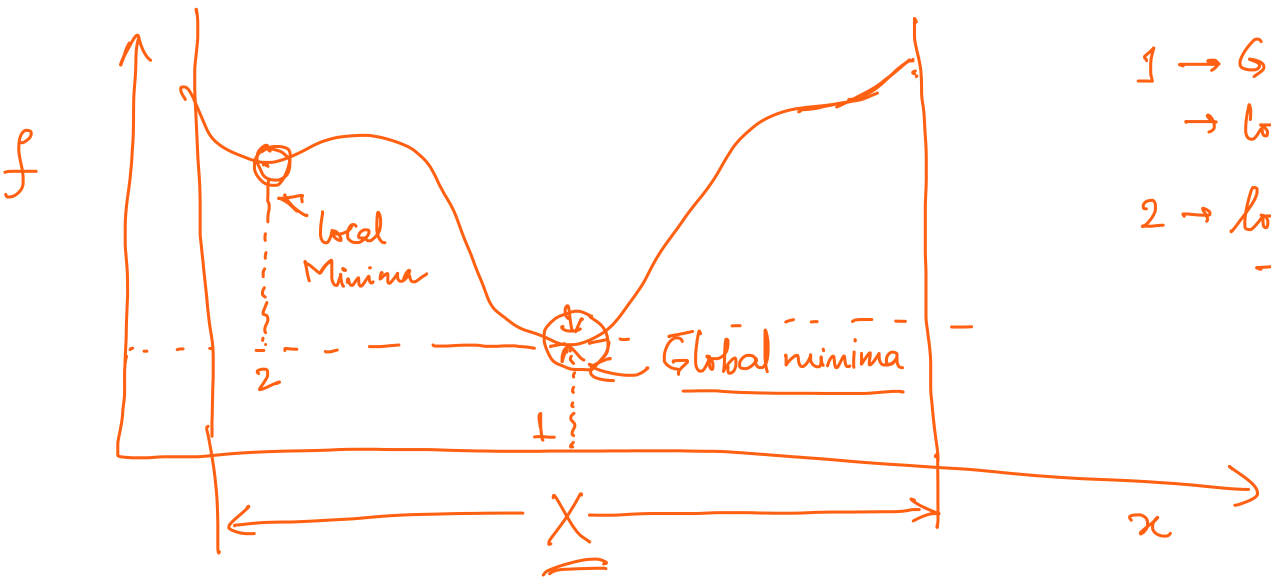
Local Minima $\rightarrow$

$$\boxed{N_\epsilon(x) = \{z \mid \|x-z\| < \epsilon\}}$$

A point $x^* \in X$ is said to be a local minima of f over X and for some $\epsilon > 0$, s.t $f(x) \geq f(x^*)$ $\forall x \in N_\epsilon(x) \cap X$.

$\leftarrow$ Neighbourhood of x.

$f(x) > f(x^*) \rightarrow$ strict local minima.



1 $\rightarrow$ Global mima.
$\rightarrow$ local mima

2 $\rightarrow$ local mima

Convex Optimization Problem

$$\min f(x)$$

s.t. $h_j(x) \leq 0 \qquad j = 1, 2, \ldots l$

$e_i(x) = 0 \qquad i = 1, 2 \ldots m$

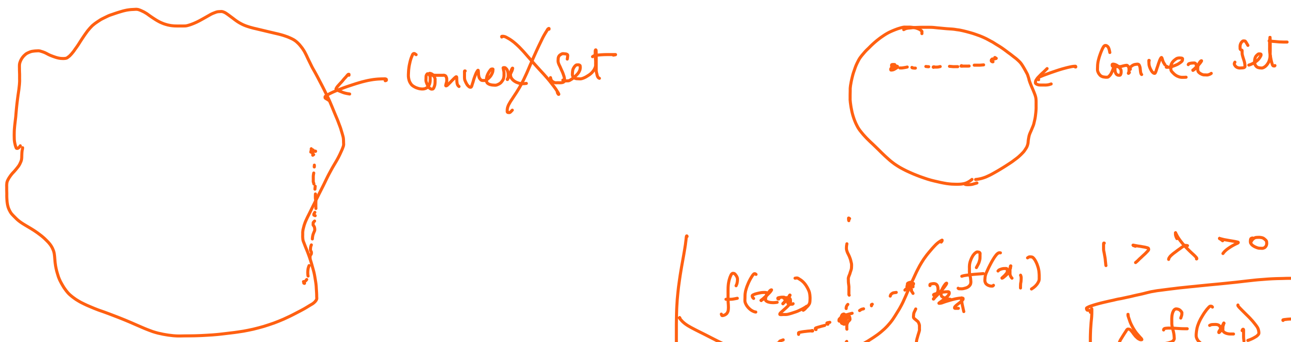
$x \in S$

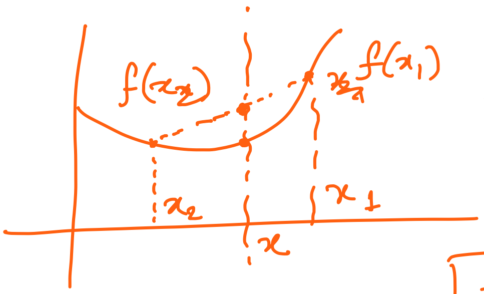
→ $f(x)$ is a convex f^n

→ $e_i(x)$ is affine i.e $e_i(x) = a_i^T x + b_i \quad \forall i \in \{1, \ldots m\}$

→ $h_j(x)$ is a convex $f^n \quad \forall j \in \{1, \ldots l\}$

→ S → Convex set.

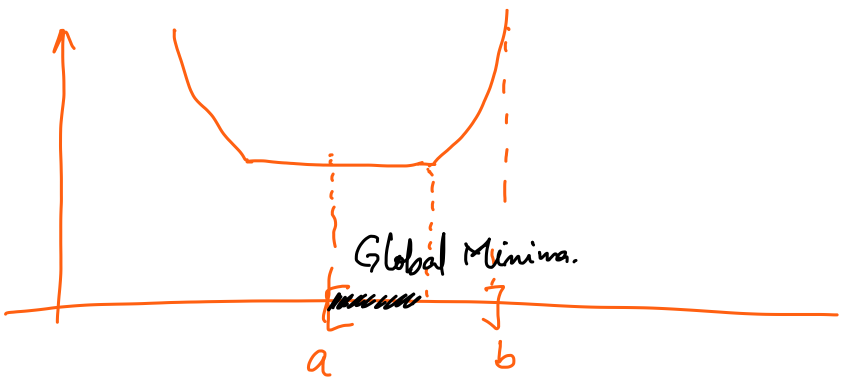




$1 > \lambda > 0$

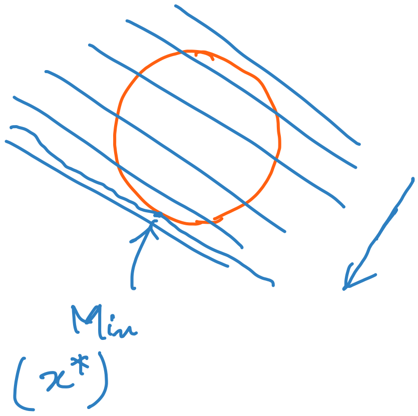
$$\boxed{\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)}$$

$\boxed{f \rightarrow \text{convex } f^n?}$

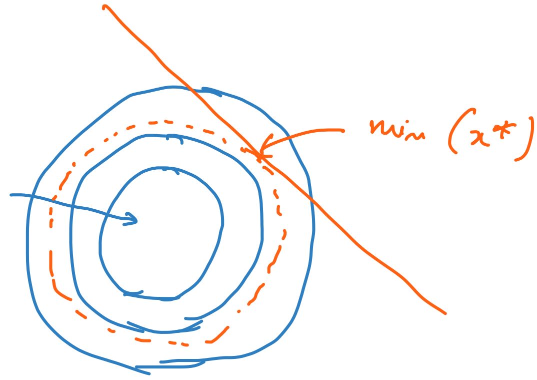


$$\min \quad x_1 + x_2$$
$$\text{s.t} \quad x_1^2 + x_2^2 = 1$$
$$x_1, x_2 \in \mathbb{R}$$

$$x_1 = \pm\sqrt{1 - x_2^2}$$



$$\min \quad x_2 \pm \sqrt{1 - x_2^2} \qquad \boxed{x_2^2 \leq 1}$$
$$x_2 \in \mathbb{R}$$



$$\min \quad x_1^2 + x_2^2$$
$$\text{s.t} \quad x_1 + x_2 = 1$$
$$x_1, x_2 \in \mathbb{R}$$

$$x_1 = 1 - x_2$$

$$\boxed{\min \quad x_2^2 + (1 - x_2)^2}$$
$$x_2 \in \mathbb{R}$$

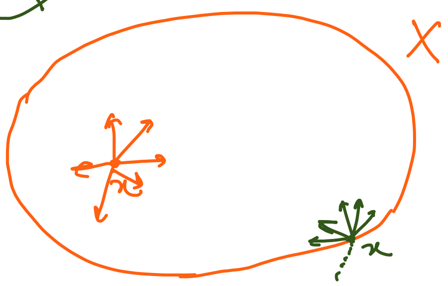
$$x_2 = \cancel{\frac{1}{\sqrt{2}}} \; \frac{1}{2} \qquad x_1 = 1/2$$

$$x^* = (1/2 \, , \, 1/2) \checkmark$$

Feasible and Descent Directions

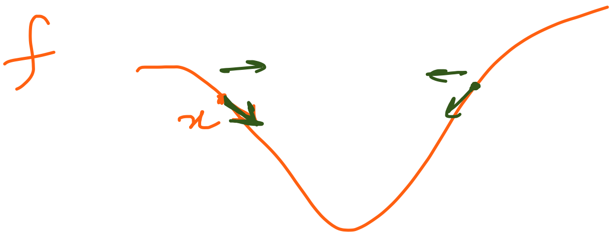
$$\min f(x)$$
$$\text{s.t. } x \in X.$$

Feasible Direction: $\vec{d} \in \mathbb{R}^n,\ d \neq 0 \longrightarrow \delta > 0$ s.t. $x + \alpha d \in X$ $\forall \alpha \in (0, \delta)$



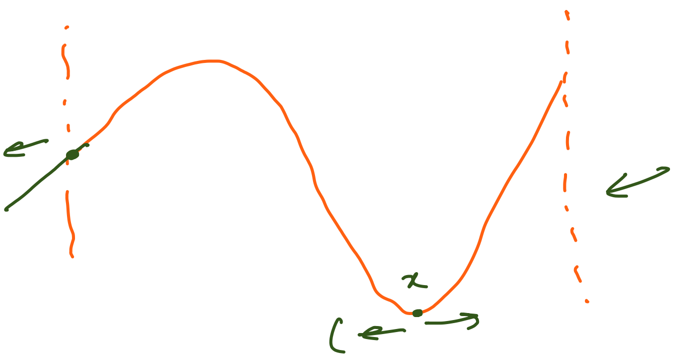
$\mathcal{F}(x) =$ Set of all feasible directions for $x \in X$.

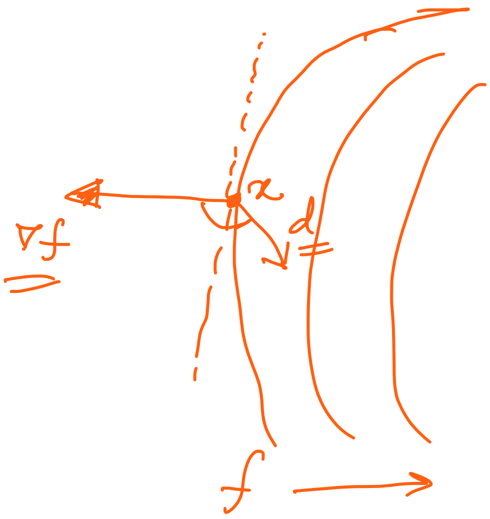
Descent Direction: $d \in \mathbb{R}^n,\ d \neq 0 \rightarrow \delta > 0$ $f(x + \alpha d) \leq f(x)$ $\forall \alpha \in (0, \delta)$



$\mathcal{D}(x) =$ Set of all descent directions at $x \in X$ (w.r.t f)

Thm: Let X be a non-empty set in $\mathbb{R}^n$ and $x^* \in X$ be a local minima of f over X. Then $\underline{\mathcal{F}(x^*) \cap \mathcal{D}(x^*) = \phi}$.





$$\{d : \nabla f(x)^T d < 0\} \leftarrow \text{Descent direction}$$

$$\bar{D}(x)$$

$$\nabla f(x)^T d < 0$$

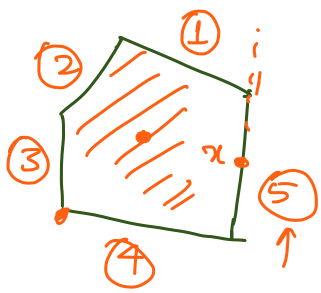
$$\lim_{\alpha \to 0^+} \frac{f(x+\alpha d) - f(x)}{\alpha} < 0$$

$$f(x+\alpha d) < f(x) \rightarrow d$$

$$\boxed{\bar{D}(x) \subseteq D(x)}$$

Consider this problem:

$$\min \; f(x)$$
$$\text{s.t} \quad h_j(x) \leq 0 \qquad j = 1, 2 \ldots l$$
$$x \in \mathbb{R}^n$$



Active Constraints :

$$A(x) = \{j : h_j(x) = 0\}$$

Lemma : For any $x \in X$,

$$\boxed{\bar{F}(x) = \{d : \nabla h_j^T(x) d < 0 \quad j \in A(x)\} \subseteq F(x)}$$

$$\boxed{\nabla h^T(x) d < 0} \leftarrow \text{Feasible Directions.}$$

By Thm 1 : $\mathcal{F}(x^*) \cap \mathcal{D}(x^*) = \phi$

$\overline{\mathcal{F}}(x^*) \cap \overline{\mathcal{D}}(x^*) = \phi$

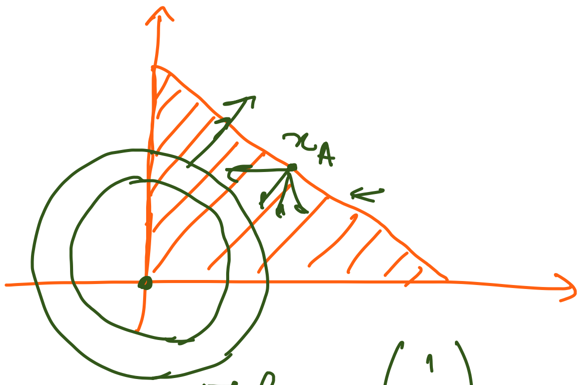
x^* is a local minima, then $\overline{\mathcal{F}}(x^*) \cap \overline{\mathcal{D}}(x^*) = \phi$

min $\underline{x_1^2 + x_2^2}$

s.t $x_1 + x_2 \le 1 \leftarrow$

$x_1 \geq 0$

$x_2 \geq 0$



$\nabla h = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$

At $x^* = (0,0) \leftarrow$

$\boxed{\overline{\mathcal{F}}(x^*) \cap \overline{\mathcal{D}}(x^*) = \phi}$

$\{ d_1 + d_2 < 0 \} \cap$

$\phi = \phi$

$\overline{\mathcal{F}}(x) = \{ d : \nabla h^T d < 0 \}$

$= \{ d_1, d_2 : d_1 + d_2 < 0 \}$

$\nabla f(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$

$\overline{\mathcal{D}}(x) = \{ d : \nabla f^T(x)\, d < 0 \}$

$= \{ d_1, d_2 : \underline{2(x_1 d_1 + x_2 d_2)} < 0 \}$

This is a necessary condition.

$$\min f(x)$$

$$\text{s.t.} \quad h_j(x) \leq 0 \qquad j = 1, \ldots, l$$

$$x \in \mathbb{R}^n$$

$$X = \{ x \in \mathbb{R}^n : h_j(x) \leq 0, \; j = 1, \ldots l \}$$

$x^* \in X$ is local minima

$$\Rightarrow \overline{F}(x^*) \cap \overline{D}(x^*) = \phi$$

$$\Rightarrow \{ d \mid \underline{\nabla h_j}^T d < 0 \quad j \in \mathcal{A}(x^*) \} \cap \{ d \mid \underline{\nabla f(x^*)}^T d < 0 \} = \phi$$

$$\underline{\underline{A}} = \begin{bmatrix} \nabla f(x^*) & \\ \vdots & \\ \nabla h_j(x^*) & j \in \mathcal{A}(x^*) \\ \vdots & \end{bmatrix}_{\underline{\underline{(1 + |\mathcal{A}(x^*)|) \times n}}}$$

$$\underline{\underline{x^*}} \in X \text{ is local minima} \Rightarrow \boxed{\{ d \mid Ad < 0 \} = \underline{\underline{\phi}}} \checkmark$$

Farka's Lemma

Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$. Then only one of the following conditions are true:

(1) $Ax \leq 0, \quad c^T x > 0$ for some $x \in \mathbb{R}^n$

(2) $A^T y = c, \quad y \geq 0$ for some $y \in \mathbb{R}^m$

Corollary of Farka's Lemma

Let $A \in \mathbb{R}^{n \times m}$. Then exactly one of the following conditions are true :-

(1) $Ax < 0$ for some $x \in \mathbb{R}^n$ $\longrightarrow$ Not True

(2) $A^T y = 0$, $y \geq 0$ for some non zero $y \in \mathbb{R}^m \leftarrow$ True

---

$x^* \in X$ is local minima $\Rightarrow \{d \mid Ad < 0\} = \phi$

$= A^T y = 0$ and $y \geq 0$

$$y = \begin{bmatrix} \lambda_0 \\ \vdots \\ \lambda_j \\ \vdots \end{bmatrix}$$

$y \neq 0$ / λ's to be non zero.

$\lambda \geq 0$

$$\boxed{\lambda_0 \nabla f(x^*) + \sum_{j \in A(x^*)} \lambda_j \nabla h_j(x^*) = 0}$$

λ_0, λ_j's are all nonzero at the same time.

Regular point: $x^* \in X$ is said to be a regular point, if gradient vectors $\nabla h_j(x^*)$ are linearly independent.

$$\boxed{x^* \in X \text{ is a regular point} \Rightarrow \lambda_0 \neq 0}$$

$$\bar{\lambda}_j = \frac{\lambda_j}{\lambda_0}$$

$$\nabla f(x^*) + \sum \bar{\lambda}_j \nabla h_j(x^*) = 0$$

$\bar{\lambda}_j \geq 0$ $\quad j \in A(x^*)$

$\boxed{\bar{\lambda}_i = 0}$ $\quad i \notin A(x^*)$

$\sum \bar{\lambda}_i \nabla h_i(x^*) = 0$

$$\nabla f(x^*) + \sum \lambda_k \nabla h_k(x^*) = 0 \qquad k \in \{1, \ldots, l\}$$

$$\lambda_k h_k(x^*) = 0$$

$$\lambda_k \geq 0$$

$k \notin A(x^*)$: $\lambda_k = 0$

$k \in A(x^*)$ $h_k = 0$

$k \notin A(x^*)$ $\lambda_k = 0$

$k \in A(x^*)$ $\lambda_k \geq 0$

Karush - Kuhn – Tucker (KKT) Conditions

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t} & h_j(x) \leq 0 & j \in \{1, \ldots, l\} \\ & & x \in \mathbb{R}^n \end{array}$$

$$X = \{x \in \mathbb{R}^n \mid h_j(x) \leq 0 \quad j \in \{1, \ldots l\}\}$$

KKT necessary Conditions (First Order): If $x^* \in X$ is a local minima and a regular point, then there exists a unique vector $\lambda^* = ((\lambda_1^*, \lambda_2^* \ldots \lambda_l^*))$ such that

$$\nabla f(x^*) + \sum_{j=1}^{l} \lambda_j^* \nabla h_j(x^*) = 0$$

$$\lambda_j^* h_j(x) = 0 \qquad \forall j \in \{1, \ldots l\}.$$

$$\lambda_j^* \geq 0$$

$$\boxed{(x^*, \lambda^*) \longrightarrow \text{KKT points}, \quad x^* \in X \quad \lambda^* \geq 0}$$

$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{l} \lambda_j h_j(x) \leftarrow$ Lagrangian function

$\nabla \mathcal{L}(x, \lambda) = 0$; λ_j : Lagrangian ~~m~~ multipliers : $\lambda_j \geq 0$;

$\boxed{\lambda_j^* h_j(x^*) = 0} \leftarrow$ Complementary slackness Condition.

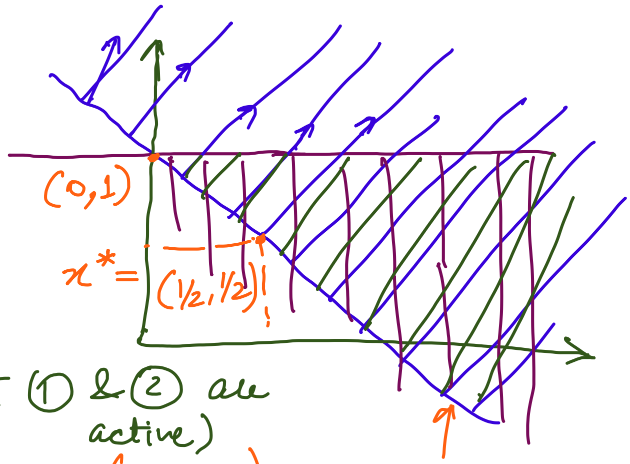
$\lambda_j^* = 0 \qquad \forall j \notin A(x^*)$

Example

$$\min \; x_1^2 + x_2^2 \qquad x_1, x_2 \in \mathbb{R}$$

$$\text{s.t} \quad x_1 + x_2 \geq 1 \quad - (1)$$

$$x_2 \leq 1 \quad - (2)$$



$$x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \lambda_1 \geq 0 \qquad \lambda_2 \geq 0$$

(Const (1) & (2) are active)

$$\mathcal{L}(x, \lambda) = x_1^2 + x_2^2 + \lambda_1 (1 - x_1 - x_2) + \lambda_2 (x_2 - 1)$$

$$\nabla \mathcal{L} = \begin{bmatrix} 2x_1 - \lambda_1 + 0 \\ 2x_2 - \lambda_1 + \lambda_2 \end{bmatrix} = 0$$

$x_1 = 0 \quad x_2 = 1$

$\lambda_1 = 0$

$\lambda_2 = -2$ X

$$x = ? \qquad \underline{\lambda_1 \geq 0} \qquad \lambda_2 = 0$$

(Const (1) is active)

$$x_1 + x_2 = 1$$

$$\mathcal{L}(x, \lambda) = x_1^2 + x_2^2 + \lambda_1 (1 - x_1 - x_2) + 0$$

$$\nabla \mathcal{L} = \begin{bmatrix} 2x_1 - \lambda_1 \\ 2x_2 - \lambda_1 \end{bmatrix} = 0$$

$\lambda_1 = 2x_1$

$\lambda_1 = 2x_2$

$\Rightarrow x_1 = x_2 = \lambda_1/2$

$$x_1 + x_2 = 1$$

$$\frac{\lambda_1}{2} + \frac{\lambda_1}{2} = 1 \Rightarrow \boxed{\lambda_1^* = 1} \geq 0$$

$$\boxed{x_1^* = x_2^* = 1/2}$$