LECTURE 19

NONLINEAR OPTIMIZATION WITH INEQUALITY CONSTRAINTS (PART I)

$$\text{(NIP)} \quad \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \\ \text{subject to} \\ \quad c_j(x) \geq 0 \quad j = 1, \dots, m \end{array}$$

NIP – nonlinear inequality-constrained program

$\boxed{\text{OBJECTIVE FUNCTION}}$ $f : \mathbb{R}^n \to \mathbb{R}$

$\boxed{\text{CONSTRAINT}}$ $c_j : \mathbb{R}^n \to \mathbb{R}$

are twice continuously differentiable

More difficult than (NEP). An (NEP) can be turned into an (NIP).

$$\begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \\ \text{subject to} \\ \quad c_j(x) = 0 \quad j = 1, \dots, m \end{array} \quad \text{equivalent to} \quad \begin{array}{l} \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \\ \text{subject to} \\ \quad \left. \begin{array}{l} c_j(x) \geq 0 \\ -c_j(x) \geq 0 \end{array} \right\} j = 1, \dots, m \end{array}$$

DEFN (Feasible Region)

The set

$$\mathbb{F} = \{x \in \mathbb{R}^n : c_j(x) \geq 0 \quad j = 1, \ldots, m\}$$

is called the feasible region.

The definitions of a local and a global minimizer are same as the definitions for (NEP) but using the feasible region as defined above.

DEFN (Active & Inactive Constraints)

Let $x_* \in \mathbb{F}$. A constraint $c_j(x)$ is said to be

* active at x_* if $c_j(x_*) = 0$, and
* inactive at x_* if $c_j(x_*) > 0$.

REMARKS

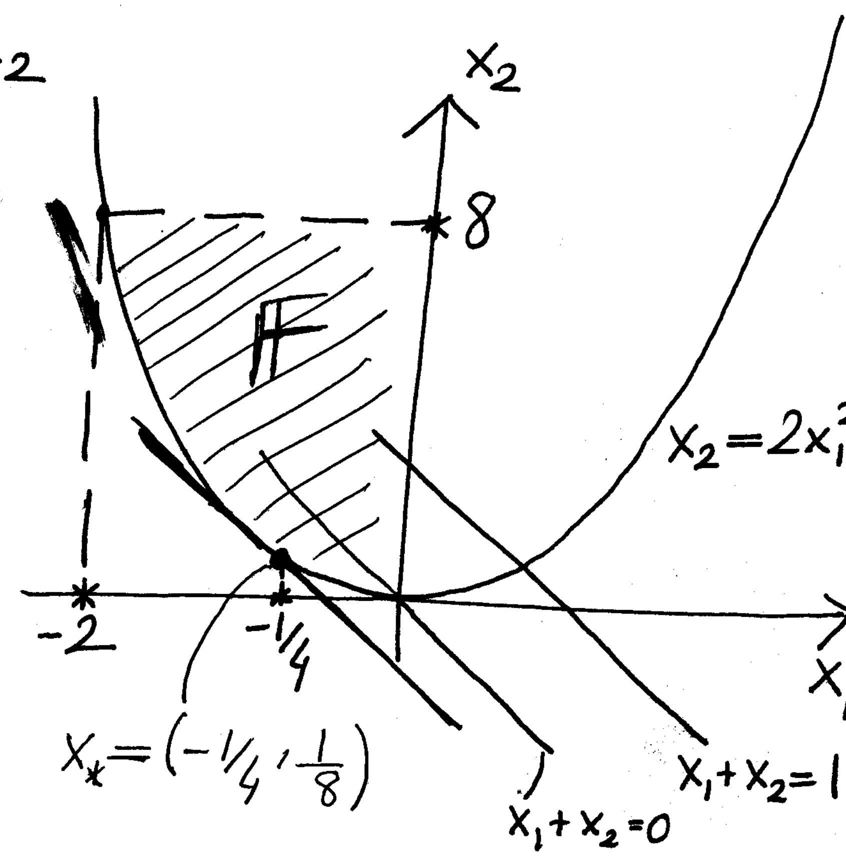
* The inherent difficulty in (NIP) is that the active constraints at a local minimizer are not known in advance.
* If the active constraints were known, an (NIP) is simply an (NEP).

EXAMPLE

$$\min_{x \in \mathbb{R}^2} \; x_1 + x_2$$

subject to

$$-2x_1^2 + x_2 \geq 0$$
$$-x_1 \geq 0$$
$$x_1 \geq -2$$
$$-x_2 \geq -8$$



At the minimizer $x_* = (-1/4, 1/8)$ only the constraint $-2x_1^2 + x_2 \geq 0$ is active.

The problem is equivalent to

$$\min_{x \in \mathbb{R}^2} \; x_1 + x_2$$

subject to

$$-2x_1^2 + x_2 = 0$$

REMARK

In general if we know precisely that constraints $j_1, j_2, \dots, j_p$ are active at a local minimizer, then (NIP) is equivalent to

$$\min_{x \in \mathbb{R}^n} \; f(x)$$

subject to

$$g_{j_k}(x) = 0 \quad k = 1, \dots, p$$

OPTIMALITY CONDITIONS

Suppose $x_* \in \mathbb{F}$ is a local minimizer.
For any feasible path $x(\alpha)$

$$f(x(\alpha)) = f(x(0)) + \left.\frac{df(x(\alpha))}{d\alpha}\right|_{\alpha=0} \alpha + \left.\frac{d^2 f(x(\alpha))}{d\alpha^2}\right|_{\alpha=t} \frac{\alpha^2}{2}$$

$$= f(x_*) + \alpha \nabla f(x(0))^T x'(0) + c\alpha^2$$

where $\alpha > 0$ and $t \in (0, \alpha)$.

Since $x_* \in \mathbb{F}$ is a local minimizer

$$0 \leqslant f(x(\alpha)) - f(x_*) = \alpha \left(\nabla f(x_*)^T x'(0) \right) + c\alpha^2$$

$$\Longrightarrow$$

$$\nabla f(x_*)^T x'(0) \geqslant 0$$

for all feasible path $x(\alpha)$ at x_*

DEFN (Feasible path for (NIP))
A twice continuously differentiable directed parametric curve $x(\alpha) : \mathbb{R} \to \mathbb{R}^n$ is a feasible path at x_* if

(i) $x(0) = x_*$,

(ii) $x'(0) \neq 0$, and

(iii) there exists a $\sigma > 0$ s.t. $c_i(x(\alpha)) \geqslant 0$ for all $\alpha \in [0, \sigma)$ and $i = 1, \cdots, m$

DEFN (Tangent Cone)

The set

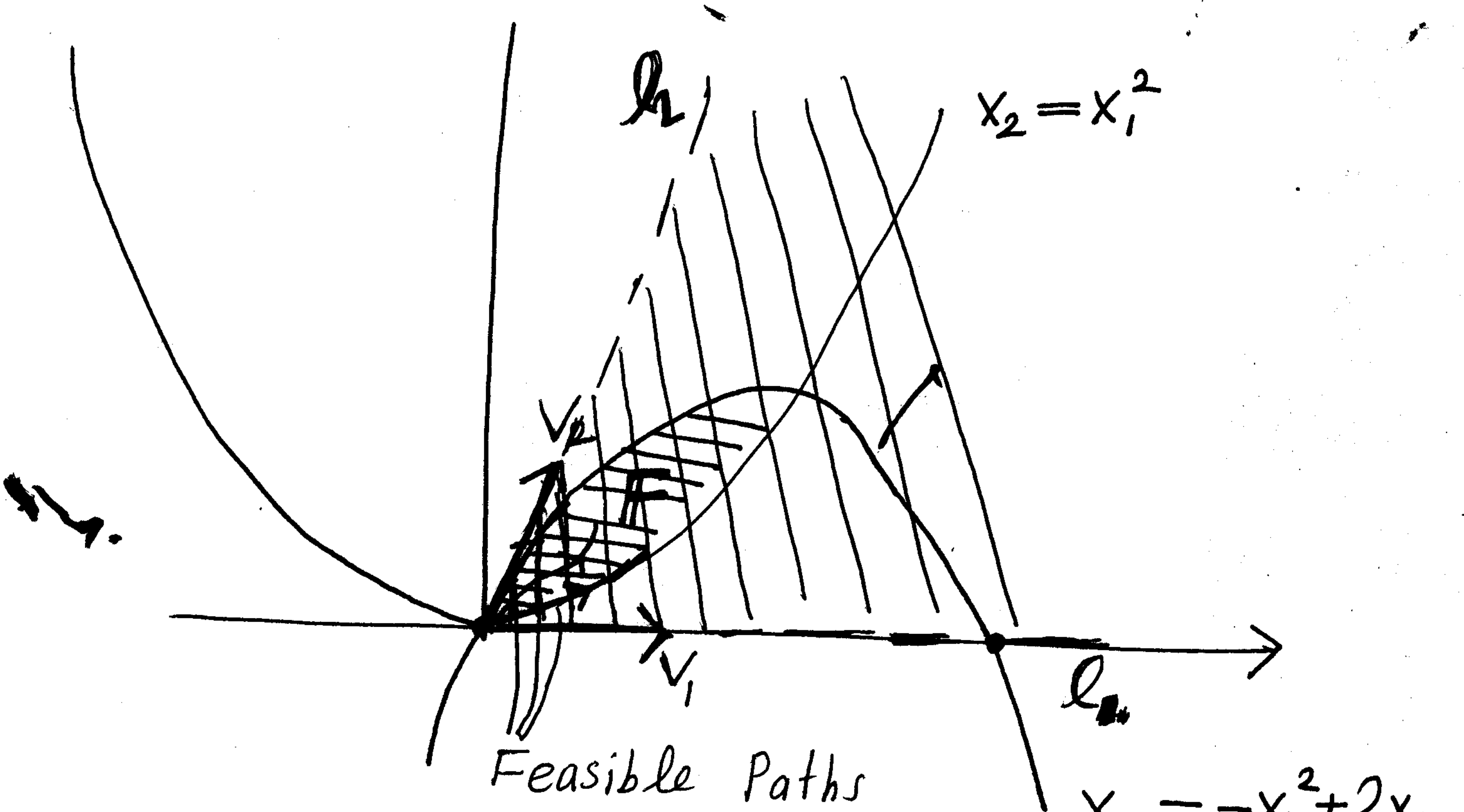
$$T^{\circ}(x_*) = \{x'(0) : x(t) \text{ is a feasible path}\} \cup \{0\}$$

is called the tangent cone at x_*.

EXAMPLE

Consider the constraints

(i) $-x_1^2 + x_2 \geq 0$ (ii) $-x_1^2 + 2x_1 - x_2 \geq 0$



V_1 – vector tangent to $x_2 = x_1^2$ at $(0,0)$

V_2 – vector tangent to $x_2 = -x_1^2 + 2x_1$ at $(0,0)$

$$T^\circ(0) = \{\alpha_1 v_1 + \alpha_2 v_2 : \alpha_1, \alpha_2 \geqslant 0\}$$

All linear combinations of v_1, v_2 with positive weights.

l_1 – line tangent to $x_2 = x_1^2$ (and parallel to v_1) at $(0,0)$

l_2 – line tangent to $x_2 = -x_1^2 + 2x_1$ (and parallel to v_2) at $(0,0)$.

$T^\circ(0)$ is the region bounded by l_1 and l_2 in the first quadrant.

$$\text{Slope of } l_2 = \left.\frac{d(-x_1^2+2x_1)}{dx_1}\right|_{x_1=0} = 2$$

$$\Longrightarrow$$

$$v_2 = \alpha_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \forall \alpha_2 \in \mathbb{R}^+$$

$$\text{Slope of } l_1 = \left.\frac{d(x_1^2)}{dx_1}\right|_{x_1=0} = 0$$

$$\Longrightarrow$$

$$v_1 = \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \forall \alpha_1 \in \mathbb{R}^+$$

$$T^{\circ}(0) = \left\{ \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} : \alpha_1, \alpha_2 \geqslant 0 \right\}$$

CHARAC 1

x_* is a local minimizer

$$\Longrightarrow$$

$\nabla f(x_*)^T p \geqslant 0$ for all $p \in T^{\circ}(x_*)$