LECTURE 5

CONSTRAINED OPTIMIZATION

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

$$c_i(x) = 0, \quad i \in E$$

$$c_i(x) \geq 0, \quad i \in I$$

E - indices of equality constraints

I - indices of inequality constraints

$f, c_i : \mathbb{R}^n \to \mathbb{R}$

twice continuously differentiable functions

Ex

① Linear programs

$$\min_{x \in \mathbb{R}^n} c^T x$$

$$Ax = b$$

$$x \geq 0 \ (\rightarrow x_j \geq 0 \quad j = 1, \ldots, n)$$

$A \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ - given

② Quadratic programs

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T H x + g^T x$$

$$Ax = b$$

$$x \geq 0$$

$H \in \mathbb{R}^{n\times n}$ symmetric, $g \in \mathbb{R}^n$ – given

$A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ – given

$x_* \in \mathbb{R}^n$ is a local minimizer of (NLP) if there exists $\delta > 0$ such that

$$f(x) \geqslant f(x_*) \quad \text{for all } x \in B(x_*, \delta) \cap F$$

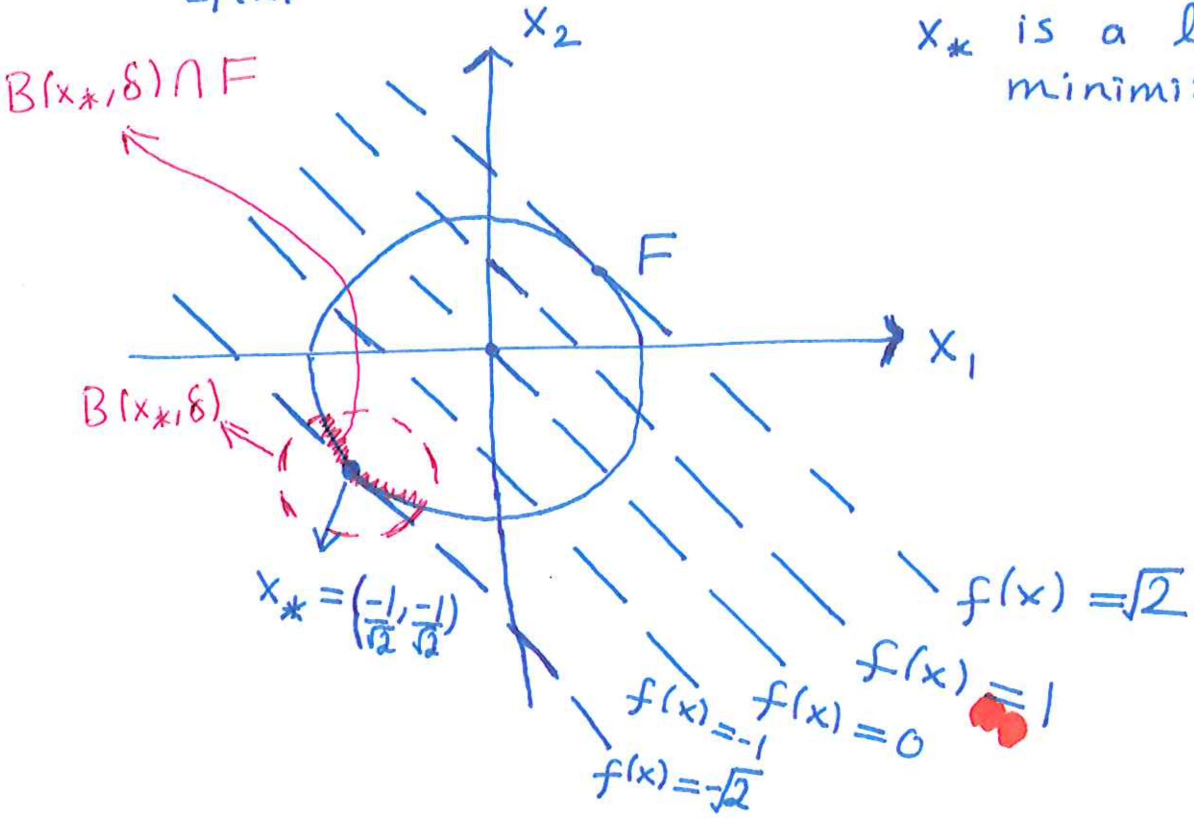
where

$$F := \{x \in \mathbb{R}^n \mid c_i(x) = 0,\ i \in E \\ c_i(x) \geqslant 0,\ i \in I\}$$

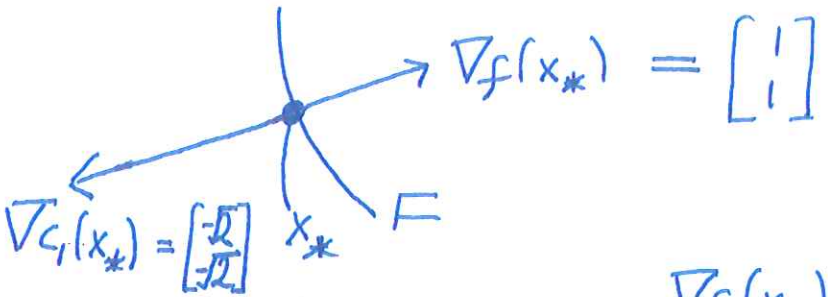
Ex

① $\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad x_1 + x_2$

$$\underbrace{-x_1^2 - x_2^2 + 1}_{c_1(x)} = 0$$



x_* is a local minimizer

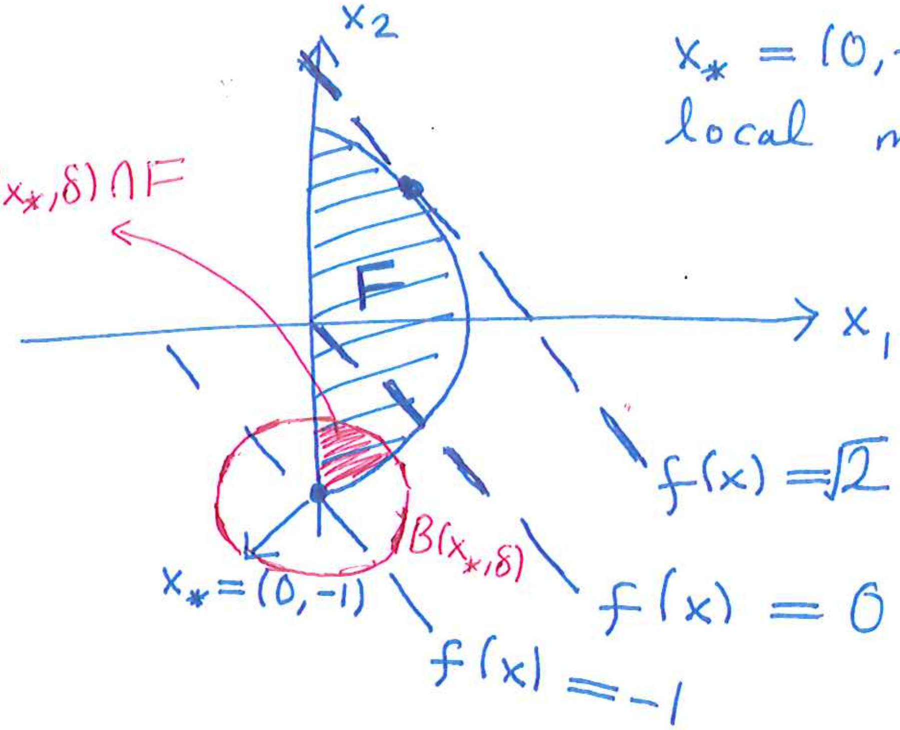


$$\nabla f(x_*) = \frac{-1}{\sqrt{2}} \nabla c_1(x_*)$$

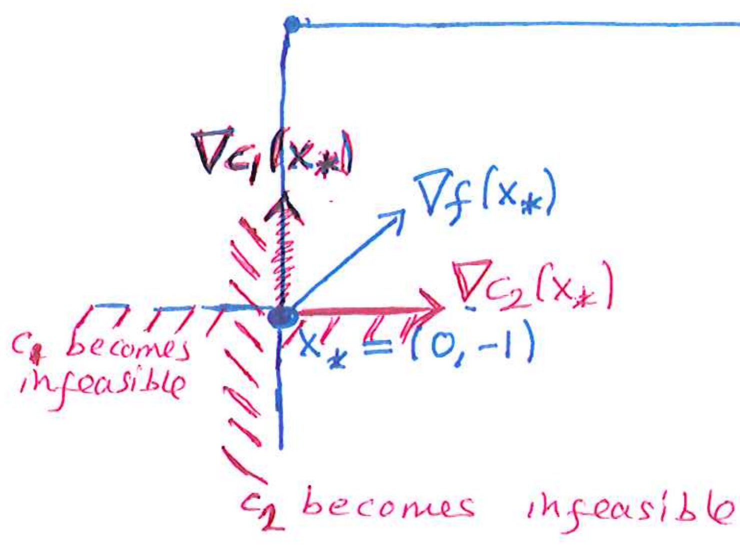
② minimize$_{x \in \mathbb{R}^2}$ $x_1 + x_2$

$c_1(x) := -x_1^2 - x_2^2 + 1 \geqslant 0$

$c_2(x) := x_1 \geqslant 0$



$x_* = (0,-1)$ is a local minimizer

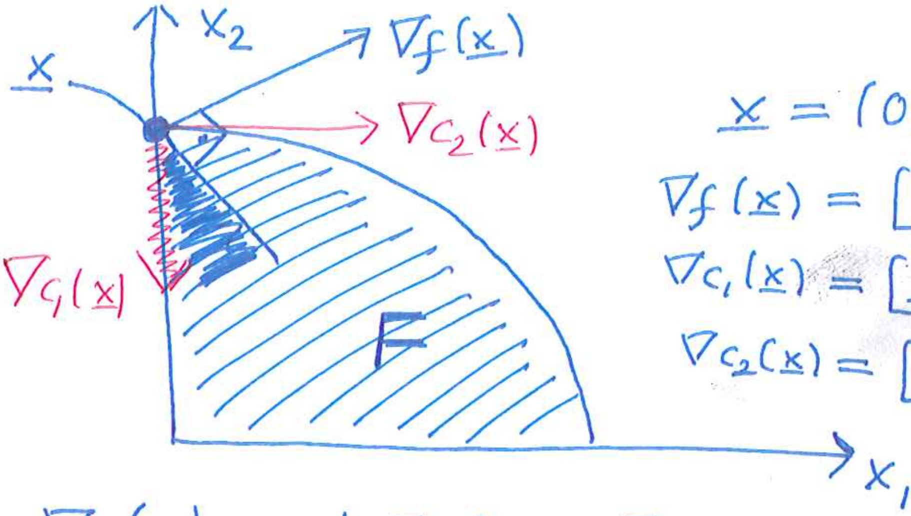


$$\nabla f(x_*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\nabla c_1(x_*) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\nabla c_2(x_*) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\nabla f(x_*) = \frac{1}{2} \nabla c_1(x_*) + \nabla c_2(x_*)$$



$\underline{x} = (0,1)$

$\nabla f(\underline{x}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\nabla c_1(\underline{x}) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$

$\nabla c_2(\underline{x}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$\nabla f(\underline{x}) = \frac{-1}{2} \nabla c_1(\underline{x}) + \nabla c_2(\underline{x})$$

yet $\underline{x}$ is not a local minimizer:

(i) heavily shaded region is in F and $f(x) < f(\underline{x})$ for all x in this region;

(ii) there exists $d \in \mathbb{R}^2$ such that

$$\nabla c_1(\underline{x})^T d > 0$$

$$\nabla c_2(\underline{x})^T d > 0$$

$$\nabla f(\underline{x})^T d < 0.$$

Derivation of first order optimality conditions

~~A path $x(\alpha) : \mathbb{R} \to \mathbb{R}^n$ is called feasible from $\underline{x}$ if~~

~~(i) $c_i(x(\alpha)) = 0, \ \forall i \in E$~~

~~$c_i(x(\alpha)) \geq 0, \ \forall i \in I$~~

~~for all $\alpha > 0$ sufficiently small.~~

~~(ii) $x(0) = \underline{x}$, and~~

~~(iii) $x'(0) \neq 0$.~~

A feasible sequence $\{x^{(k)}\}$ (leading) to $x_* \in F$ is such that

(i) $x^{(k)} \in F$ for all k large enough,

(ii) $\lim_{k \to \infty} x^{(k)} = x_*$ and

(iii) $x^{(k)} \neq x_*$ for all k.

Ex

$$\begin{aligned} \underset{x \in \mathbb{R}^2}{\text{minimize}} \quad & x_1^2 + x_2^2 \\ -x_1^2 - x_2^2 + 1 &\geq 0 \\ x_1 &\geq 0 \end{aligned}$$

some feasible sequences to $(0,1)$

$$\left\{\begin{bmatrix} 1/k \\ \sqrt{1-1/k^2} \end{bmatrix}\right\} \quad \left\{\begin{bmatrix} 0 \\ 1-1/k \end{bmatrix}\right\} \quad \left\{\begin{bmatrix} 1/k \\ 1-1/k \end{bmatrix}\right\}$$

Let $\{x^{(k)}\}$ be a feasible sequence to x_*. If any subsequence $\{x^{(k_j)}\}$ of $\{x^{(k)}\}$ is such that

$$\lim_{j \to \infty} \frac{x^{(k_j)} - x_*}{\| x^{(k_j)} - x_* \|} = d,$$

then d is called a limiting direction at x_*.

Regarding the example above

$$\left\{\begin{bmatrix} 1/k \\ \sqrt{1-1/k^2} \end{bmatrix}\right\} \quad \left\{\begin{bmatrix} 0 \\ 1-1/k \end{bmatrix}\right\} \quad \left\{\begin{bmatrix} 1/k \\ 1-1/k \end{bmatrix}\right\}$$

$$d = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad d = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad d = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

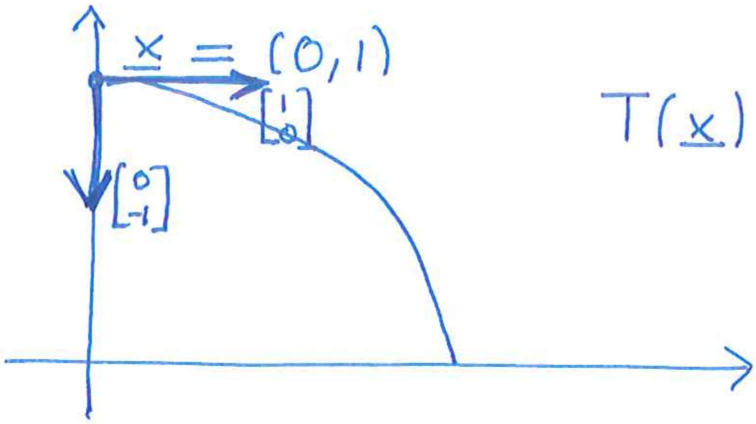
Tangent cone

$$T(x_*) = \{ \alpha d \mid \alpha \in \mathbb{R}^+ \text{ and } d \text{ is a limiting direction at } x_* \}$$

Ex

Constraints

$$-x_1^2 - x_2^2 + 1 \geq 0 \quad \text{and} \quad x_1 \geq 0$$



$$T(\underline{x}) = \left\{ \alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \;\middle|\; \alpha_1, \alpha_2 \in \mathbb{R}^+ \right\}$$

$$= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;\middle|\; x_1 \geq 0, \; x_2 \leq 0 \right\}$$

LEMMA

x_* is a local minimizer for (NLP)

$$\Longrightarrow$$

$$\nabla f(x_*)^T d \geq 0 \quad \text{for all } d \in T(x_*).$$

PROOF

Suppose

$$\nabla f(x_*)^T d < 0$$

for some $d \in T(x_*)$. Without loss of generality, suppose $\|d\|_2 = 1$ (i.e. d is a limiting direction at x_*).

Let $\{x^{(k)}\}$ be a feasible sequence to x_* such that

$$\lim_{k\to\infty} \frac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} = d.$$

But then, by Taylor's thm (with 2nd order remainder)

$$f(x^{(k)}) = f(x_*) + \nabla f(x_*)^T d \,\|x^{(k)} - x_*\| + O(\|x^{(k)} - x_*\|^2)$$

$$< f(x_*) + \frac{1}{2} \nabla f(x_*)^T d \,\|x^{(k)} - x_*\|$$

$$\Longrightarrow \quad f(x^{(k)}) < f(x_*)$$

for each k sufficiently large.

Since $\lim_{k\to\infty} x^{(k)} = x_*$, x_* is not a local minimizer. $\square$

Algebraic characterization of the tangent cone

Indices of active constraints at x_*

$$A(x_*) := E \cup \{ j \in I \mid c_j(x_*) = 0 \}$$

Ex

$$\begin{aligned} &\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad x_1 + x_2 \\ &-x_1^2 - x_2^2 + 1 \geqslant 0 \\ &\quad x_1 \geqslant 0 \end{aligned}$$

$$A\left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) = \{1\}$$

$$A(0,1) = \{1,2\}$$

$$A(0,0) = \{2\}$$

Let d be a limiting direction at x_*, and $\{x^{(k)}\}$ be a feasible sequence to x_* such that

$$\lim_{k \to \infty} \frac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} = d.$$

1 If $j \in E$,

$$\begin{aligned} 0 &= c_j(x^{(k)}) \\ &= c_j(x_*) + \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\| \\ &\qquad + O(\|x^{(k)} - x_*\|^2) \end{aligned}$$

for all k large enough. Suppose $\nabla c_j(x_*)^T d \neq 0$, for instance suppose

$\nabla c_j(x_*)^T d > 0$ (argument similar to the one below applies if $\nabla c_j(x_*)^T d < 0$). But then

$$c_j(x^{(k)}) > c_j(x_*) + \frac{1}{2} \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\|$$

$$= \frac{1}{2} \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\| > 0$$

We obtain the contradiction that $x^{(k)}$ is not feasible for all large k. Consequently,

$$\nabla c_j(x_*)^T d = 0.$$

2 If $j \in \mathcal{I} \cap \mathcal{A}(x_*)$,

$$0 \leq c_j(x^{(k)})$$

$$= c_j(x_*) + \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\| + O(\|x^{(k)} - x_*\|^2)$$

$$= \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\| + O(\|x^{(k)} - x_*\|^2)$$

for all k large enough. Suppose $\nabla c_j(x_*)^T d < 0$. But then

$$c_j(x^{(k)}) < \frac{1}{2} \nabla c_j(x_*)^T d \, \|x^{(k)} - x_*\| < 0$$

which contradicts with the assumption that $x^{(k)}$ is feasible for large k. Consequently,

$$\nabla c_j(x_*)^T d \geq 0$$

In terms of Jacobians of constraints

$$(\text{Equality}) \quad c_E(x) := \begin{bmatrix} c_{j_1}(x) \\ \vdots \\ c_{j_p}(x) \end{bmatrix}, \quad E = \{j_1, j_2, \ldots, j_p\}$$

$$J_E(x) = c_E'(x) = \begin{bmatrix} \nabla c_{j_1}(x)^T \\ \vdots \\ \nabla c_{j_p}(x)^T \end{bmatrix}$$

$$\begin{matrix} \nabla c_j(x)^T d = 0 \\ j \in E \end{matrix} \iff J_E(x)\, d = 0$$

$$(\text{Inequality}) \quad c_{IA}(x) := \begin{bmatrix} c_{k_1}(x) \\ \vdots \\ c_{k_\ell}(x) \end{bmatrix}, \quad I \cap A(x) = \{k_1, \ldots, k_\ell\}$$

$$J_{IA}(x) = c_{IA}'(x)$$

$$\begin{matrix} \nabla c_j(x)^T d \geq 0 \\ j \in I \cap A(x) \end{matrix} \iff J_{IA}(x)\, d \geq 0$$

<u>THM</u>

$$\begin{aligned} T(x_*) &\subseteq \{ d \mid \nabla c_j(x_*)^T d = 0 \ \ j \in E, \\ &\qquad \nabla c_j(x_*)^T d \geq 0 \ \ j \in I \cap A(x_*)\} \\ &= \{ d \mid J_E(x_*)\, d = 0 \text{ and } \\ &\qquad J_{IA}(x_*)\, d \geq 0 \} \end{aligned}$$

Ex

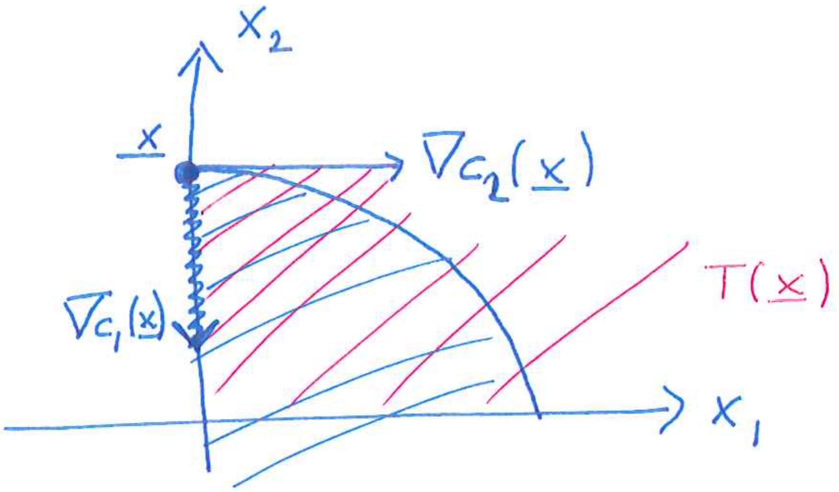
Constraints

$$-x_1^2 - x_2^2 + 1 \geqslant 0 \quad \text{and} \quad x_1 \geqslant 0$$

let $\underline{x} = (0, 1)$.

$$T(\underline{x}) = \left\{ \alpha_1 \begin{bmatrix} 0 \\ -2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \;\middle|\; \alpha_1, \alpha_2 \in \mathbb{R}^+ \right\}$$

$$= \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;\middle|\; x_1 \geqslant 0 \text{ and } x_2 \leqslant 0 \right\}$$



$$S = \{ d \mid J_{IA}(\underline{x})\, d \geqslant 0 \}$$

$$= \left\{ d \;\middle|\; \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix} d \geqslant 0 \right\}$$

$$= \left\{ \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \;\middle|\; d_1 \geqslant 0 \text{ and } d_2 \leqslant 0 \right\}$$

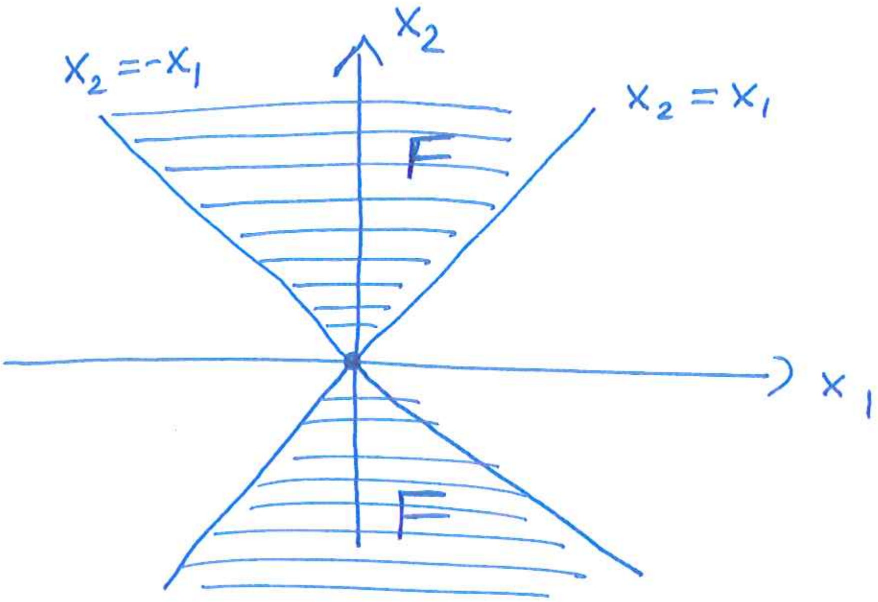
$$= T(\underline{x})$$

Ex

Constraint

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

let $x_* = (0, 0)$.



$$T(x_*) = \left\{ \alpha_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \;\middle|\; \begin{array}{c} \alpha_1, \alpha_2 \geq 0 \\ \text{OR} \\ \alpha_1, \alpha_2 \leq 0 \end{array} \right\}$$

$$= F$$

$$S = \{ d \mid \nabla c_1(x_*)^T d \geq 0 \}$$

$$= \{ d \mid [0 \quad 0]\, d \geq 0 \}$$

$$= \mathbb{R}^2$$

Thus

$$T(x_*) \subset \mathbb{R}^2$$

We say that the constraint qualification holds at x_* if

$$T(x_*) = \{ d \mid J_E(x_*)d = 0 \text{ and } J_{IA}(x_*)d \geq 0 \}$$

THM (LICQ - Linear Independence Constraint Qualification)

Suppose that $\{ \nabla c_j(x_*) \mid j \in A(x_*) \}$ is linearly independent. Constraint qualification holds at x_*.

PROOF

Let $\{t_k\}$ be a positive sequence in $\mathbb{R}$ such that $\lim_{k \to \infty} t_k = 0$. It suffices to show that for any $d \in \mathbb{R}^n$ such that $\|d\|_2 = 1$ and

$$\nabla c_j^T d = 0 \qquad j \in E,$$
$$\nabla c_j^T d \geq 0 \qquad j \in I \cap A(x_*),$$

we have $d \in T(x_*)$. Without loss of generality assume all inequality constraints are active, i.e., $A(x_*) = E \cup I$.

~~Consider the map G~~

Let $\underset{m \times n}{J(x)} = c'(x)$ and $Z \in \mathbb{C}^{n \times (n-m)}$ whose columns form an orthonormal basis for $\text{Null}(J(x_*))$ so that

$$J(x_*)\,Z = 0.$$

Define the map $G : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ by

$$G(x,t) := \begin{bmatrix} c(x) - t\,J(x_*)\,d \\ Z^T(x - x_* - t\,d) \end{bmatrix}.$$

Since

$$G_x(x_*, 0) = \underbrace{\begin{bmatrix} J(x_*) \\ Z^T \end{bmatrix}}_{n \times n}$$

is invertible, by the implicit function thm, there exists an open neighborhood N of 0 and a continuous function $x(t) : N \to \mathbb{R}^n$ such that $x(0) = x_*$ and

$$G(x(t), t) = 0 \quad \text{for all } t \in N.$$

Define $x^{(k)} := x(t_k)$ for large k. We claim $\{x^{(k)}\}$ is a feasible sequence to x_*. For each large k, we have

$$c(x^{(k)}) - t_k J(x_*)\,d = 0$$

$$\Longrightarrow \underset{j \in E}{0 = c_j(x^{(k)}) - t_k \nabla c_j(x_*)^T d = c_j(x^{(k)})}$$

$$\underset{j \in I}{0 = c_j(x^{(k)}) - t_k \nabla c_j(x_*)^T d \leqslant c_j(x^{(k)})}$$

Consequently, $x^{(k)} \in F$. Furthermore,

$$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} x(t_k) = x\left(\lim_{k\to\infty} t_k\right) = x(0) = x_*.$$

Furthermore, assuming $x^{(k)} = x_*$ for some k leads to

$$0 = \begin{bmatrix} c(x^{(k)}) - t_k J(x_*) d \\ Z^T(x^{(k)} - x_* - t_k d) \end{bmatrix}$$

$$= \begin{bmatrix} -t_k J(x_*) d \\ -t_k Z^T d \end{bmatrix} = \begin{bmatrix} J(x_*) \\ Z^T \end{bmatrix} (-t_k d)$$

which is a contradiction, because $\begin{bmatrix} J(x_*) \\ Z^T \end{bmatrix}$ is invertible, $t_k \neq 0$ and $d \neq 0$. Thus, $x^{(k)} \neq x_*$ for each large k.

Finally, we show

$$\lim_{k\to\infty} \frac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} = d,$$

implying $d \in T(x_*)$. By Taylor's thm

$$c(x^{(k)}) - t_k J(x_*) d = 0$$

$$\cancel{c(x_*)} + J(x_*)(x^{(k)} - x_*) - t_k J(x_*) d + O(\|x^{(k)} - x_*\|^2) = 0$$

Dividing both sides by $\|x^{(k)} - x_*\|$, and taking the limit as $k \to \infty$ yield

$$\lim_{k \to \infty} \begin{bmatrix} J(x_*) \\ Z^T \end{bmatrix} \frac{\left[(x^{(k)} - x_*) - t_k d \right]}{\|x^{(k)} - x_*\|} = 0$$

(recall $\begin{bmatrix} J(x_*) \\ Z^T \end{bmatrix}$ is invertible)

$$\Longrightarrow \lim_{k \to \infty} \frac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} - \frac{t_k}{\|x^{(k)} - x_*\|} d = 0$$

$$\Longrightarrow \lim_{k \to \infty} \frac{t_k}{\|x^{(k)} - x_*\|} = 1 \quad \text{and}$$

$$\lim_{k \to \infty} \frac{x^{(k)} - x_*}{\|x^{(k)} - x_*\|} = d,$$

completing the proof. $\square$

Implicit function theorem

<u>THM</u>

Suppose $F: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a continuously differentiable function, $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ are such that

$$F_z(z,x)\Big|_{z=a,\ x=b}$$

is invertible and $F(a,b) = 0$.

There exists an open neighborhood N of b and a continuously differentiable function $\phi: N \to \mathbb{R}^n$ such that

(i) $\phi(b) = a$,

(ii) $F(\phi(x), x) = 0 \quad \forall x \in N$.

Simple case

$\cancel{f: \mathbb{R} \to \mathbb{R}}$ $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

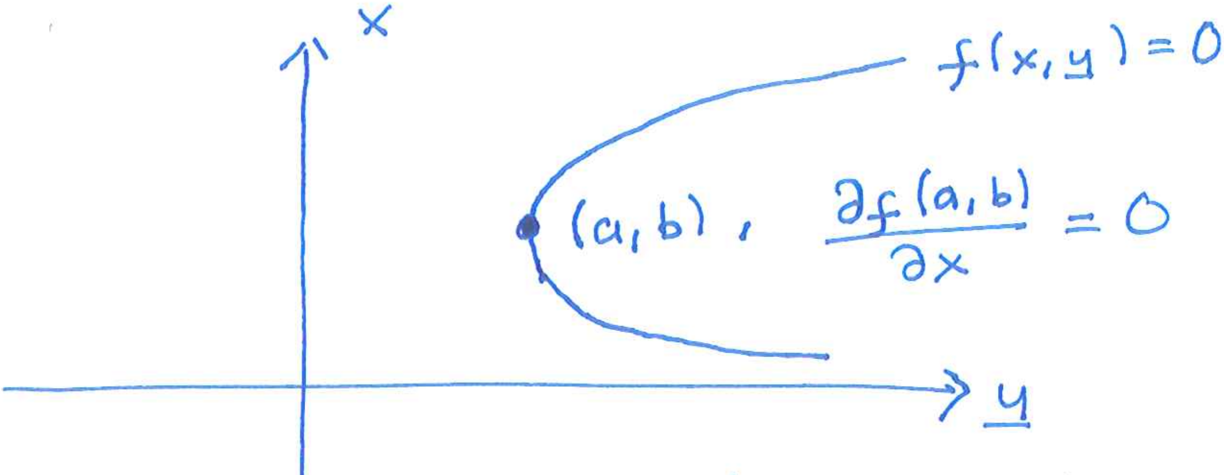
continuously differentiable,

$\partial f(x,y)/\partial x \Big|_{\substack{x=a\\ y=b}} \neq 0$ and $f(a,b) = 0$

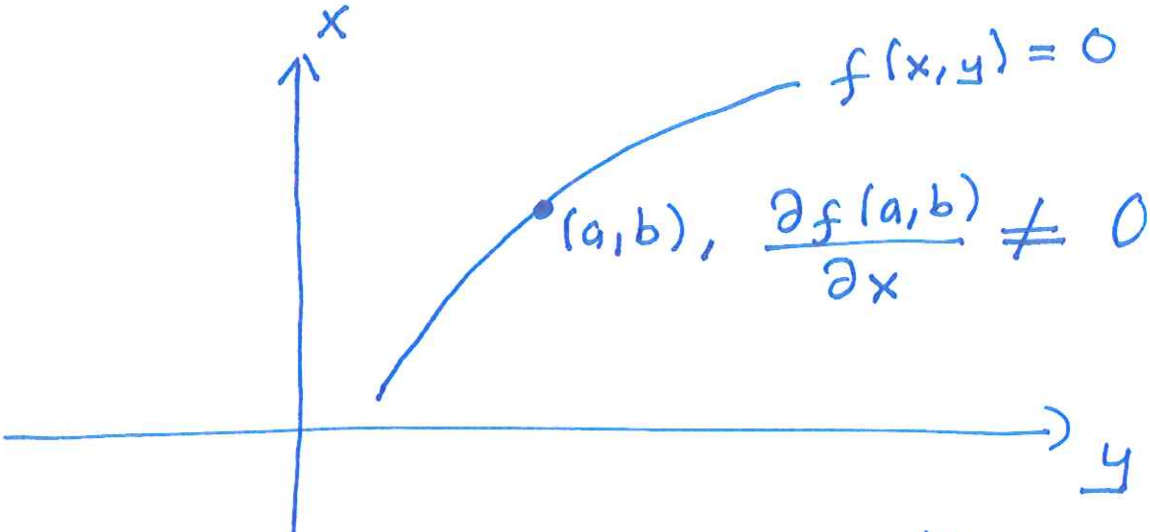
$\exists N$ of b and $\phi: N \to \mathbb{R}$

$$f(\phi(y), y) = 0 \quad \forall y \in N$$

$$\phi(b) = a$$



x - not a function of y near (a,b)



x - a function of y near (a,b)

First Order Necessary Condition

THM

The following are equivalent:

(1) $\nabla f(x_*)^T d \geq 0$ for all $d \in \mathbb{R}^n$ such that

(+) $\nabla c_j(x_*)^T d \geq 0 \quad j \in I \cap A(x_*)$,

(++) $\nabla c_j(x_*)^T d = 0 \quad j \in E$;

(2) $$\nabla f(x_*) = \sum_{j \in A(x_*)} \lambda_j \nabla c_j(x_*)$$

for some $\lambda_j \in \mathbb{R}$ such that $\lambda_j \geq 0$ for each $j \in I \cap A(x_*)$.

PROOF

(2) $\Longrightarrow$ (1)

Suppose (2) holds. For each $d \in \mathbb{R}^n$ satisfying (+) and (++), we have

$$\nabla f(x_*)^T d = \sum_{j \in E} \lambda_j \underbrace{\nabla c_j(x_*)^T d}_{=0} + \sum_{j \in I \cap A(x_*)} \lambda_j \underbrace{\nabla c_j(x_*)^T d}_{\geq 0}$$

$$\geq 0.$$

$\sim(2) \implies \sim(1)$

Let us define

$$N := \left\{ \sum_{j \in A(x_*)} \lambda_j \nabla c_j(x_*) \;\middle|\; \lambda_j \in \mathbb{R}, \quad \lambda_j \geq 0 \quad j \in I \cap A(x_*) \right\}.$$

Suppose (2) does not hold, that is $\nabla f(x_*) \notin N$. N is a closed convex cone, so let $s_* \in N$ be such that

$$\| s_* - \nabla f(x_*) \| = \underset{s \in N}{\text{minimize}} \; \| s - \nabla f(x_*) \|.$$

We claim that

$$d := s_* - \nabla f(x_*)$$

is such that $\nabla f(x_*)^T d < 0$ and (+), (++) hold.

$\boxed{\nabla f(x_*)^T d < 0}$

Notice that

$$\| s_* - \nabla f(x_*) \| = \underset{t \geq 0}{\text{minimize}} \; \| t s_* - \nabla f(x_*) \|$$

Consequently,

$$\left. \frac{d \; \| t s_* - \nabla f(x_*) \|}{dt} \right|_{t=1} = 0$$

$$\Longrightarrow \left.\frac{2t s_*^T s_* - 2\nabla f(x_*)^T s_*}{\| t s_* - \nabla f(x_*) \|}\right|_{t=1} = 0$$

$$\Longrightarrow (x) \quad s_*^T \underbrace{(s_* - \nabla f(x_*))}_{d} = 0 .$$

It follows that

$$\nabla f(x_*)^T d = (s_* - d)^T d$$
$$= s_*^T d - \|d\|^2$$
$$= -\|d\|^2 < 0 .$$

$$\boxed{\begin{array}{ll} \nabla c_j(x_*)^T d = 0 & j \in E \\ \nabla c_j(x_*)^T d \geqslant 0 & j \in I \cap A(x_*) \end{array}}$$

For each $s \in N$, we have

$$\| s_* - \nabla f(x_*) \| = \underset{\theta \in [0,1]}{\text{minimize}} \| s_* + \theta(s - s_*) - \nabla f(x_*) \|$$

that is

$$\| s_* + \theta(s - s_*) - \nabla f(x_*) \|^2 \geqslant \| s_* - \nabla f(x_*) \|^2$$
$$\forall \theta \in [0,1]$$

$$\Longrightarrow \theta^2 \| s - s_* \|^2 + 2\theta (s - s_*)^T (s_* - \nabla f(x_*)) \geqslant 0$$
$$\forall \theta \in [0,1]$$

$$\Longrightarrow (s - s_*)^T (s_* - \nabla f(x_*)) \geqslant 0$$

$$\xRightarrow{\text{from } (x)} s^T d \geqslant 0 .$$

For each $j \in E$, $\nabla c_j(x_*) \in N$ and $-\nabla c_j(x_*) \in N$ (i.e., choose $\lambda_j = 1$ and $\lambda_j = -1$, respectively, and $\lambda_k = 0$ for each $k \in A(x_*)$ such that $k \neq j$).
Therefore,

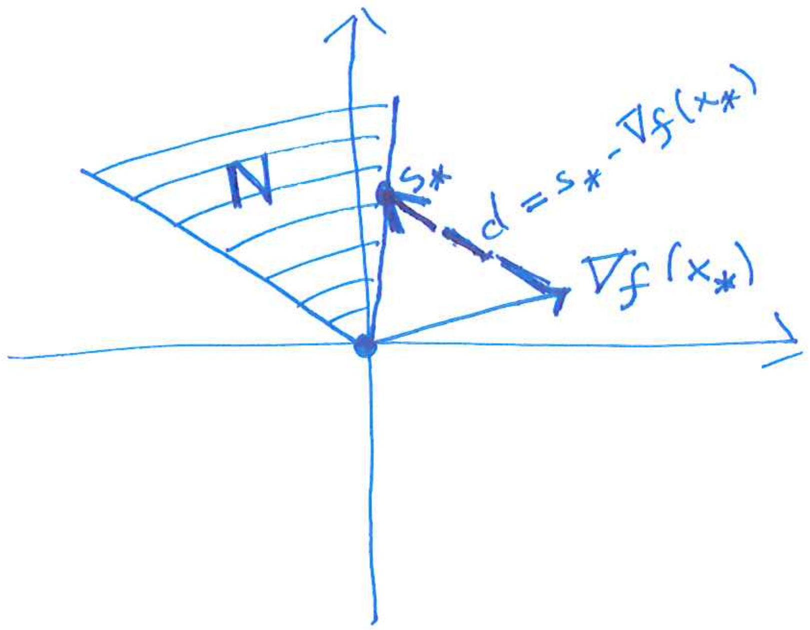
$$\nabla c_j(x_*)^T d \geq 0 \quad \text{and} \quad (-\nabla c_j(x_*))^T d \geq 0$$

$$\Longrightarrow \quad \nabla c_j(x_*)^T d = 0$$

Also, for each $j \in I \cap A(x_*)$, $\nabla c_j(x_*) \in N$ (i.e., choose $\lambda_j = 1$ and $\lambda_k = 0$ for each $k \in A(x_*)$ such that $k \neq j$). Therefore

$$\nabla c_j(x_*)^T d \geq 0.$$

We conclude that (1) does not hold.
$\square$



THM

Suppose linear independence constraint qualification holds at $x_* \in \mathbb{R}^n$. If x_* is a local minimizer of (NLP), then there exist $\lambda_j \in \mathbb{R}$ for each $j \in A(x_*)$ such that $\lambda_j \geqslant 0$ if $j \in I \cap A(x_*)$ and

$$\nabla f(x_*) = \sum_{j \in A(x_*)} \lambda_j \nabla c_j(x_*).$$

Ex

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad -2x_1 + x_2 \qquad\qquad x_* = ([redacted], 0)$$

$$(1-x_1)^3 - x_2 \geqslant 0$$

$$x_2 + 0.25x_1^2 - 1 \geqslant 0$$

$$\nabla f(x_*) = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \nabla c_1(x_*) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \nabla c_2(x_*) = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$$

$$A(x_*) = \{1, 2\}$$

$$\nabla f(x_*) = \lambda_1 \nabla c_1(x_*) + \lambda_2 \nabla c_2(x_*)$$

$$\Longrightarrow$$

$$\lambda_2 = -4 \quad \text{and} \quad \lambda_1 = -5$$

By first order necessary conditions, x_* is not a local minimizer.