LECTURE 21

NONLINEAR OPTIMIZATION WITH INEQUALITY CONSTRAINTS (PART II)

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

FIRST ORDER OPTIMALITY CONDITION FOR NIP

Suppose x_* is a point where constraint qualification holds.

x_* is a local minimizer of (NIP)

$$\Longrightarrow$$

(i) $c_j(x_*) \geqslant 0 \quad j=1,\dots,m$

(ii) $\nabla f(x_*) = J_a(x_*)^T \lambda$ for some $\lambda \geqslant 0$

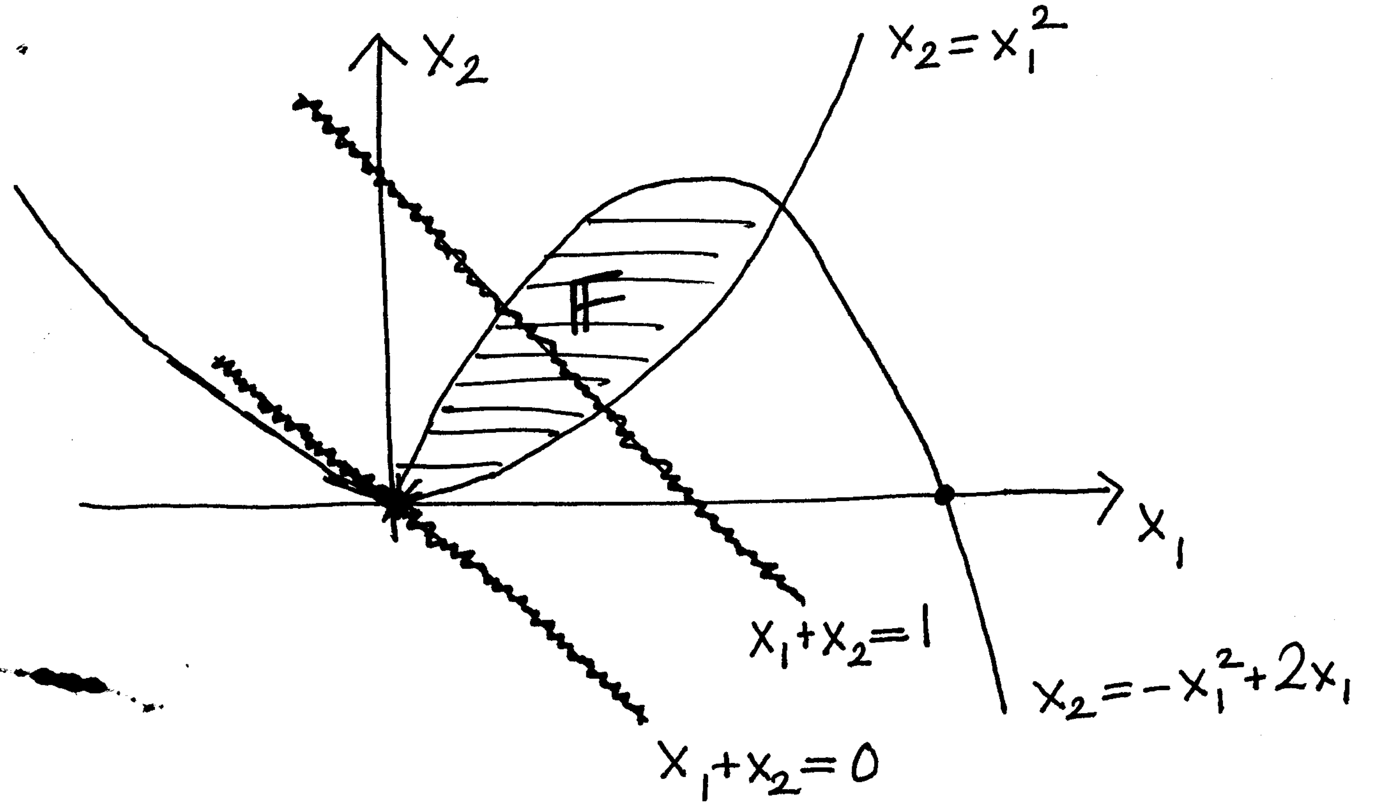
(equivalently $\nabla f(x_*) \in \mathcal{N}_a(x_*)$)

EXAMPLES

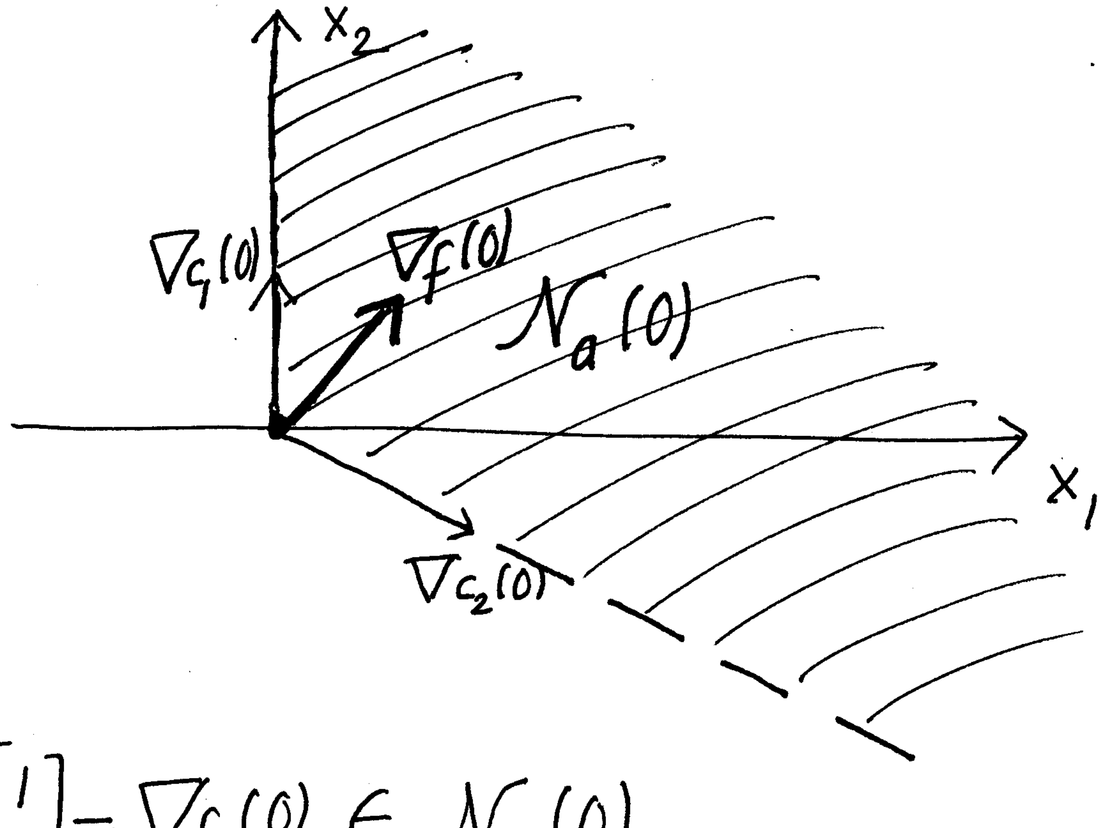
① minimize $\underbrace{x_1 + x_2}_{f(x)}$

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

$\underbrace{-x_1^2 + 2x_1 - x_2}_{c_2(x)} \geqslant 0$



$$\mathcal{N}_a(0) = \left\{ \alpha_0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} : \alpha_0, \alpha_1 \geq 0 \right\}$$



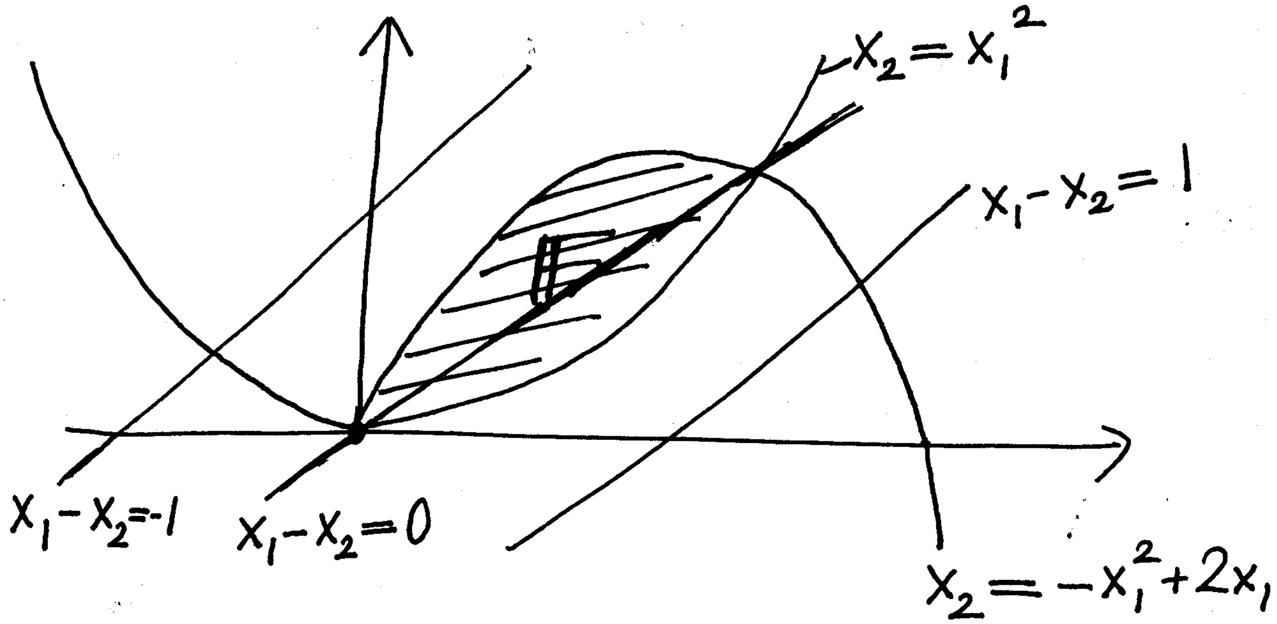
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \nabla f(0) \in \mathcal{N}_a(0)$$

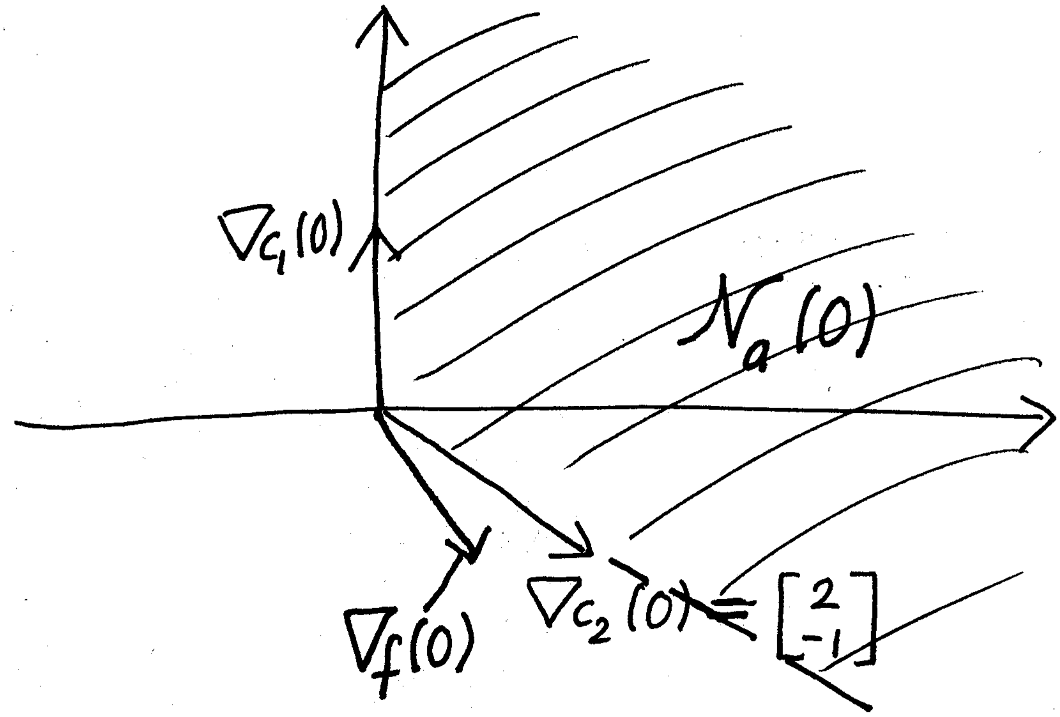
* $(0,0)$ is a local minimizer

② minimize $x_1 - x_2$

$-x_1^2 + x_2 \geq 0$

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





$$\nabla f(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \notin \mathcal{N}_a(0)$$

* (0,0) is not a local minimizer.

FARKA'S LEMMA

Following conditions are equivalent.

(1) $\nabla f(x_*)^T p \geqslant 0$ for all p s.t. $J_a(x_*)p \geqslant$

(2) $\nabla f(x_*) = J_a(x_*)^T \lambda$ for some $\lambda \geqslant 0$

PROOF OF (2) $\Longrightarrow$ (1)

Assume $\nabla f(x_*) = J_a(x_*)^T \lambda$ for some $\lambda \geqslant 0$

For all p satisfying $J_a(x_*)p \geqslant 0$

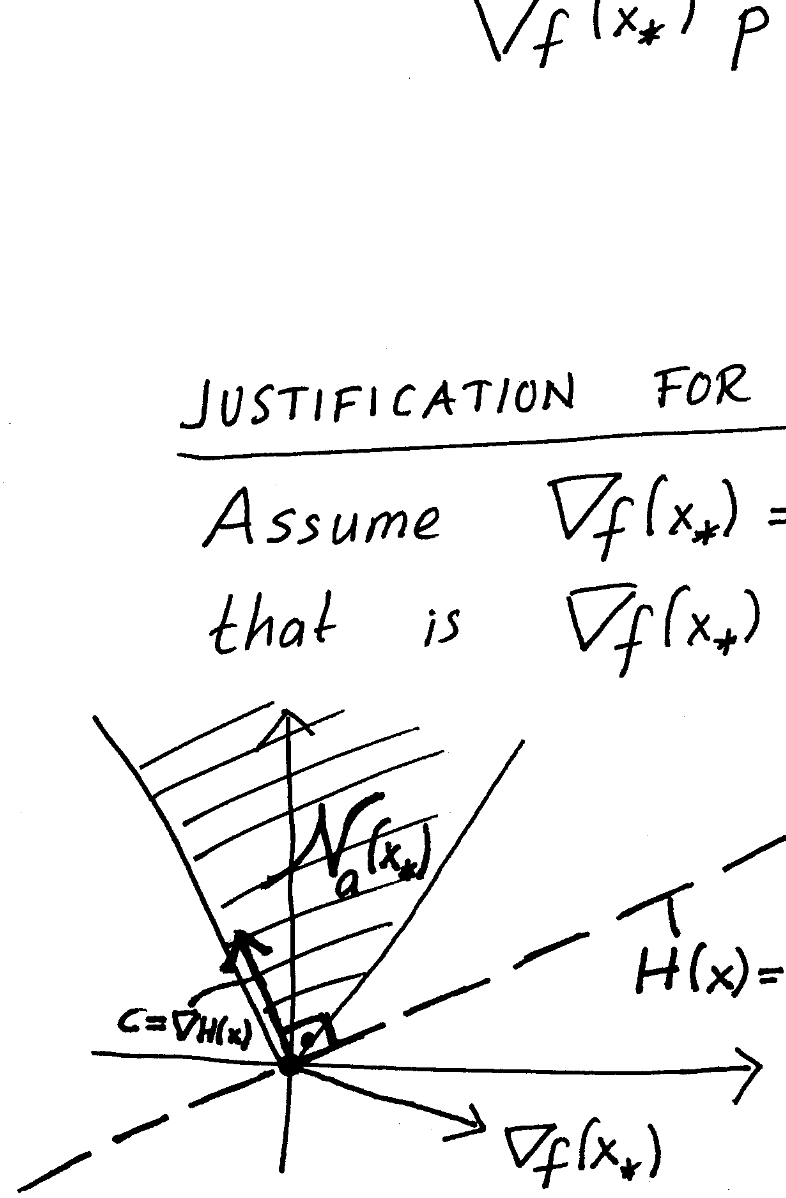
$$\nabla f(x_*)^T p = \left(J_a(x_*)^T \lambda\right)^T p$$

$$= \underbrace{\lambda^T}_{\geqslant 0} \underbrace{J_a(x_*)\,p}_{\geqslant 0} \geqslant 0$$

JUSTIFICATION FOR $\sim(2) \Longrightarrow \sim(1)$

Assume $\nabla f(x_*) \neq J_a(x_*)^T \lambda$ for all $\lambda \geqslant 0$,

that is $\nabla f(x_*) \notin \mathcal{N}_a(x_*)$.



There exists a separating hyperplane

$$H(x) = c^T x = 0$$

such that

(1) $\mathcal{N}_a(x_*) \subseteq H_c^+$

and

(2) $\nabla f(x_*) \in H_c^-$

where

$$H_c^+ = \{p : c^T p \geq 0\}$$
$$H_c^- = \{p : c^T p < 0\}$$

Consequently

(i) $\mathcal{N}_a(x_*) \subseteq H_c^+ \Longrightarrow \begin{pmatrix} \nabla c_{j_1}(x_*) \in H_c^+ \\ \vdots \\ \nabla c_{j_p}(x_*) \in H_c^+ \end{pmatrix}$

$$\Longrightarrow \begin{pmatrix} \nabla c_{j_1}(x_*)^T c \geq 0 \\ \vdots \\ \nabla c_{j_p}(x_*)^T c \geq 0 \end{pmatrix}$$

$$\Longrightarrow J_a(x_*) c \geq 0$$

(ii) $\nabla f(x_*) \in H_c^- \Longrightarrow \nabla f(x_*)^T c < 0$

There exists a feasible descent direction c such that

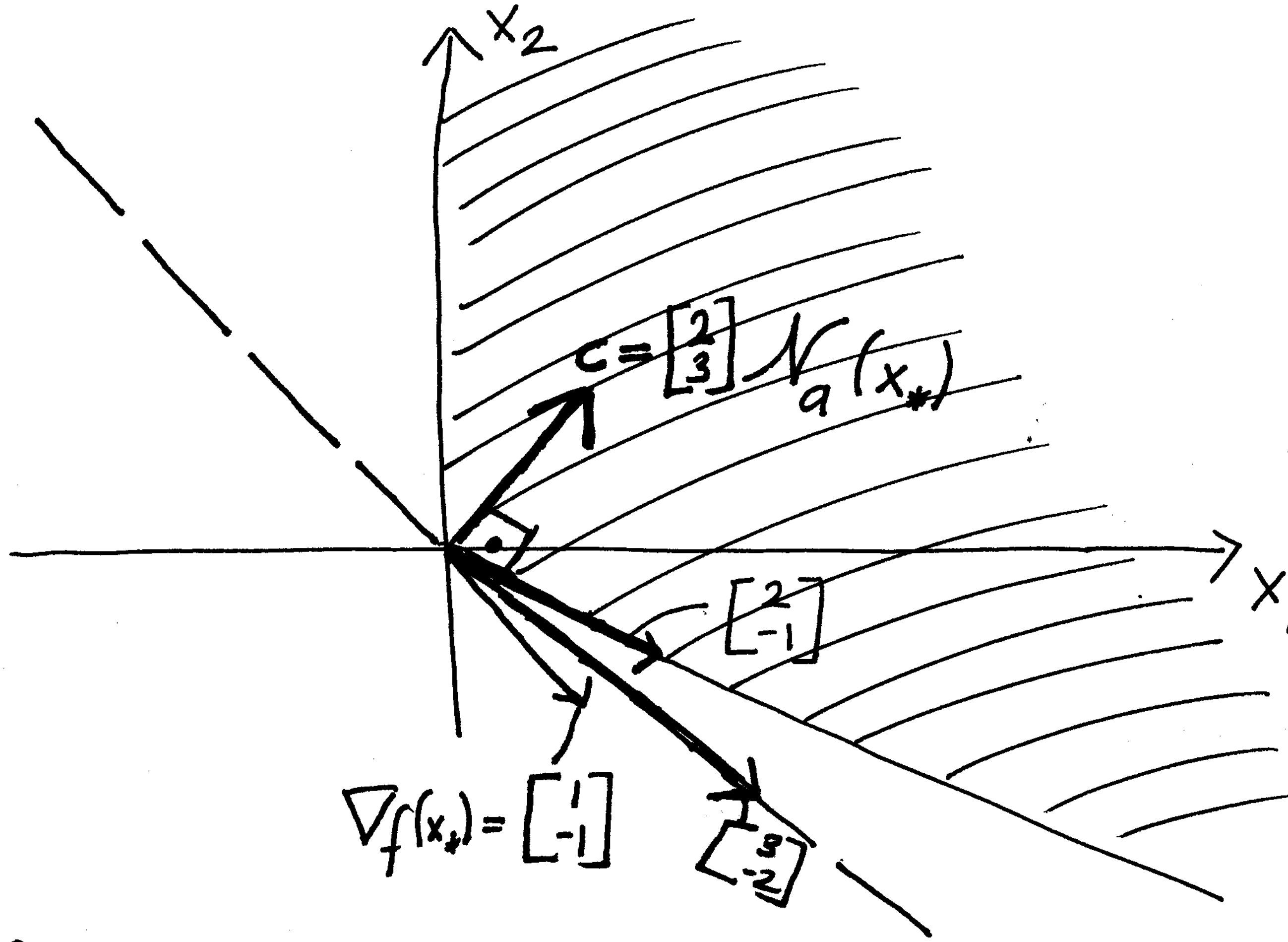
* $\nabla f(x_*)^T c < 0$ (c is descent)
* $J_a(x_*) c \geq 0$ (c is feasible).

EXAMPLE

$$\text{minimize} \quad x_1 - x_2$$

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

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



Note that $c = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is a feasible descent direction

Separating hyperplane

$$* \; \underbrace{\nabla f(0)^T c}_{\boxed{\text{DESCENT}}} = [1 \quad -1] \begin{bmatrix} 2 \\ 3 \end{bmatrix} = -1 < 0$$

$$H(x) = [2 \quad 3] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} =$$

$$* \; \underbrace{J_a(0)\, c}_{\boxed{\text{FEASIBLE}}} = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \geqslant 0$$

COMPLEMENTARITY FORM

It is more convenient to express

$$\nabla f(x) = J_a(x)^T \lambda = \lambda_1 \nabla c_{j_1}(x) + \ldots + \lambda_\ell \nabla c_{j_\ell}(x)$$

in terms of all constraints. (rather than active constraints.)

$$J_a(x)^T \lambda = \nabla f(x) \text{ for some } \lambda \geq 0$$

$$\Longleftrightarrow$$

$$J(x)^T \tilde{\lambda} = \nabla f(x) \text{ for some } \tilde{\lambda} \geq 0$$

and

$\boxed{\text{COMPLEMENTARITY CONDITION}}$ $c_j(x)\, \tilde{\lambda}_j = 0$ $\left(\begin{array}{l}\text{if } \bar{c}_j(x) \neq 0, \\ \text{then } \tilde{\lambda}_j = 0\end{array}\right)$

<u>THM</u> (Optimality Conditions in Complementarity form)

Suppose x_* is a point where constraint qualification holds. If x_* is a local minimizer, then there exists $\lambda \in \mathbb{C}^m$ s.t.

(i) $c_j(x_*) \geq 0 \quad j = 1, \ldots, m$

(ii) $\nabla f(x_*) = J(x_*)^T \lambda$

(iii) $\lambda_j \geq 0 \quad j = 1, \ldots, m$

(iv) $c_j(x_*)\, \lambda_j = 0 \quad j = 1, \ldots, m$

We say that the <u>strict complementarity condition</u> holds at x_* if

$*$ $c_j(x_*)=0$ OR $\lambda_j=0$

BUT NOT $c_j(x_*)=\lambda_j=0$

for $j=1,\dots,m$

GENERAL NONLINEAR OPTIMIZATION

$$\text{(NP)}\qquad \begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)\\ &\text{subject to}\\ &\quad c_j(x)\geqslant 0 \quad j=1,\dots,m\\ &\quad \tilde{c}_i(x)=0 \quad i=1,\dots,p \end{aligned}$$

NP – Nonlinear program

(as usual f, c_j, $\tilde{c}_i$ are twice continuously differentiable)

(NP) can be expressed as

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)\\ &\text{subject to}\\ &\quad c_j(x)\geqslant 0\\ &\quad \tilde{c}_i(x)\geqslant 0\\ &\quad -\tilde{c}_i(x)\geqslant 0 \end{aligned}$$

LINEAR INDEPENDENCE CONSTRAINT QUALIFICATION

It can be shown that the constraint qualification holds at x_* if

$$\{\nabla c_{j_1}(x_*), \dots, \nabla c_{j_\ell}(x_*), \nabla \tilde{c}_1(x_*), \dots, \nabla \tilde{c}_p(x_*)$$

is linearly independent where

$$\mathcal{A}(x_*) = \{j_1, \dots, j_\ell\}$$

is the set of active inequality constraints.

Assuming constraint qualification holds at $x_* \in \mathbb{F}$.

x_* is a local minimizer

$$\Longrightarrow$$

$$\nabla f(x_*) = \sum_{j=1}^{m} \lambda_j \nabla c_j(x_*)$$

$$+$$

$$\sum_{i=1}^{\ell} (\lambda_i^+ - \lambda_i^-) \nabla \tilde{c}_i(x_*)$$

for some $\lambda_j, \lambda_i^+, \lambda_i^- \geq 0$
(and complementarity condition holds)

$$\Longleftrightarrow$$

$$\nabla f(x_*) = \sum_{j=1}^{m} \lambda_j \nabla c_j(x_*) + \sum_{i=1}^{\ell} \tilde{\lambda}_i \nabla \tilde{c}_i(x_*)$$

for some $\lambda_j \geq 0$ and $\tilde{\lambda}_i$.
(and complementarity condition holds)

$$\Longleftrightarrow$$

$$\nabla f(x_*) = J(x_*)^T \lambda + \tilde{J}(x_*)^T \tilde{\lambda}$$

where $\lambda \geqslant 0$. (and complementarity condition holds)

NOTATION

Above and elsewhere

$$J(x) = c'(x) = \begin{bmatrix} \nabla c_1(x)^T \\ \vdots \\ \nabla c_m(x)^T \end{bmatrix}$$

$$\tilde{J}(x) = \tilde{c}'(x) = \begin{bmatrix} \nabla \tilde{c}_1(x)^T \\ \vdots \\ \nabla \tilde{c}_\ell(x)^T \end{bmatrix},$$

-KKT-

THM (Karush-Kuhn-Tucker Conditions for NP)

Let x_* be a point where the constraint qualification holds. If x_* is a local minimizer of (NP), then there exists $\lambda \in \mathbb{R}^m$ and $\tilde{\lambda} \in \mathbb{R}^\ell$ such that

(i) $c_j(x_*) \geqslant 0 \quad j=1,\dots,m$

(ii) $\tilde{c}_i(x_*) = 0 \quad i=1,\dots,\ell$

(iii) $\nabla f(x_*) = J(x_*)^T\lambda + \tilde{J}(x_*)^T\tilde{\lambda}$

(iv) $\lambda_j \geqslant 0 \quad j=1,\dots,m$

(v) $c_j(x_*)\lambda_j = 0 \quad j=1,\dots,m$

EXAMPLES

① Linear Program

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad c^T x$$

subject to

$$Ax = b$$
$$x \geq 0$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

Constraint qualification always holds.

KKT conditions

If x_* is a local minimizer, there exists $\lambda \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^m$ such that

(i) $Ax_* = b$

(ii) $x_* \geq 0$

(iii) $c = \underbrace{A^T \mu}_{\text{equality}} + \underbrace{\lambda}_{\text{inequality}}$

(iv) $\lambda \geq 0$

(v) $\lambda_j (x_*)_j = 0 \quad j = 1, \ldots, m$

② Quadratic Program

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \frac{1}{2} x^T H x + g^T x$$

subject to

$$Ax \geqslant b$$

$$\tilde{A} x = \tilde{b}$$

where $A \in \mathbb{R}^{m \times n}$, $\tilde{A} \in \mathbb{R}^{l \times n}$, $b \in \mathbb{R}^n$, $\tilde{b} \in \mathbb{R}$

Constraint qualification holds.

<u>KKT conditions</u>

If x_* is a local minimizer, there exist $\lambda \in \mathbb{R}^m$ and $\mu \in \mathbb{R}^l$ such that

(i) $\tilde{A} x_* = \tilde{b}$

(ii) $A x_* \geqslant b$

(iii) $H x_* + g = \tilde{A}^T \mu + A^T \lambda$

(iv) $\lambda \geqslant 0$

(v) $\lambda_j (A x_* - b)_j = 0 \quad j = 1, \dots, m$