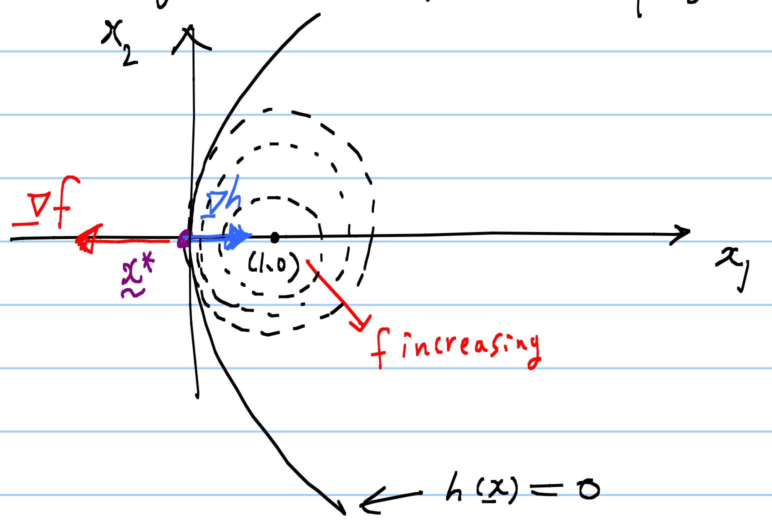
Lecture 4-2 : NLP for Equality Constraints

노트 제목

Example : $\text{Min } f(\underset{\sim}{x}) = (x_1-1)^2 + x_2^2 + 1$

subject to $h(\underset{\sim}{x}) = x_1 - \frac{1}{4}x_2^2 = 0$



▣ Solution: $\underset{\sim}{x}^* = (0,0)^T$

1) By inspection using graphs

OR 2) Using the reduced objective function

$x_1 - \frac{1}{4}x_2^2 = 0$

$\boxed{x_2 = t\,;\ x_1 = \frac{1}{4}t^2} \Rightarrow \tilde{f}(t) \equiv f(\frac{1}{4}t^2, t)$

$= \frac{1}{16}t^4 + \frac{1}{2}t^2 + 2$

→ Solve unconstrained problem in "t"

$\frac{d\tilde{f}}{dt} = \frac{1}{4}t^3 + t = 0 \rightarrow t^* = 0$

$\Rightarrow \underline{x}^* = (0,0)$

▣ Observation at $\underset{\sim}{x} = \underset{\sim}{x}^*$

$\underline{\nabla} f(\underline{x}^*) \ /\!/ \ \underline{\nabla} h(\underline{x}^*)$

General Constrained Problems with 2 Design Variables

$$\min f(x_1, x_2)$$
$$\text{st} \quad h(x_1, x_2) = 0$$

STEP 1 : Rewrite $h(x_1, x_2) = 0$ as

$$\left.\begin{aligned} x_1 &= \xi_1(t) \\ x_2 &= \xi_2(t) \end{aligned}\right\} \text{Only one Indepedent variable}$$

STEP 2 : Define $\tilde{f}(t) = f(\underbrace{\xi_1(t)}_{\to x_1}, \underbrace{\xi_2(t)}_{\to x_2})$

and solve

$$\min \tilde{f}(t) \qquad \text{(Unconstrain)}$$

At optimal point $\underset{\sim}{x}^*$ (i.e, t^*)

$$\left.\frac{d\tilde{f}}{dt}\right|_{t=t^*} = 0$$

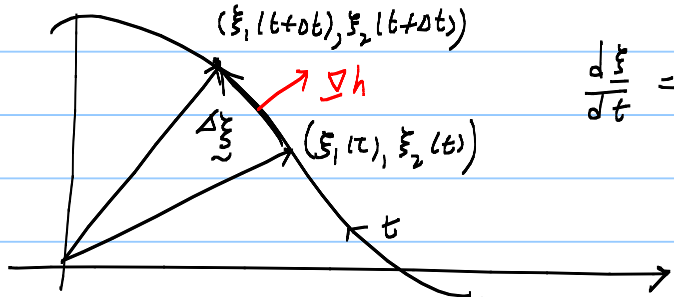
Thus

$$0 = \frac{d\tilde{f}}{dt} = \frac{\partial f}{\partial x_1}\frac{d\xi_1}{dt} + \frac{\partial f}{\partial x_2}\frac{d\xi_2}{dt}$$

$$= \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right) \cdot \left(\frac{d\xi_1(t)}{dt}, \frac{d\xi_2(t)}{dt}\right)$$

ie $$\boxed{\nabla f^T(x^*) \begin{Bmatrix} \dfrac{d\xi_1(t^*)}{dt} \\ \dfrac{d\xi_2(t^*)}{dt} \end{Bmatrix} = 0}$$

$\Rightarrow$



$$\frac{d\xi}{dt} = \lim_{\Delta t \to 0} \frac{\xi(t+\Delta t) - \xi(t)}{\Delta t}$$

$h = 0 \Leftrightarrow x_1 = \xi_1(t)$ & $x_2 = \xi_2(t)$

ⓐ $\nabla f \perp \dfrac{d\xi}{dt}$ at $x = x^*$ $(t = t^*)$

↳ tangent vector

ⓑ Note : $h(x_1, x_2) = h(\xi_1(t), \xi_2(t)) = 0$

$$0 = \frac{dh}{dt} = \frac{\partial h}{\partial x_1}\frac{\partial \xi_1}{\partial t} + \frac{\partial h}{\partial x_2}\frac{\partial \xi_2}{\partial t} \Leftrightarrow \boxed{\nabla h^T \frac{d\xi}{dt} = 0}$$

$$\boxed{\nabla f(x^*) + \lambda \nabla h(x^*) = 0 \quad \text{at } x = x^*}$$

Necessary Conditions (Lagrange Condition)

if $\underset{\sim}{x}^*$ is a Local min pt of $f(\underset{\sim}{x})$
subject to $h_i(\underset{\sim}{x}) = 0 \quad (i=1, \cdots, \ell)$
and CQ (Constraint Qualification) holds at $\underset{\sim}{x}^*$,

then there exists a vector $\underset{\sim}{\lambda} = (\lambda_1, \lambda_2, \cdots \lambda_\ell)^T$
such that

$$\underset{\sim}{\nabla} f(\underset{\sim}{x}^*) + \sum_{i=1}^{\ell} \lambda_i \nabla h_i(\underset{\sim}{x}^*) = 0$$

where λ_i are called the Lagrange multipliers.

The Method of Lagrange ~~***~~

1. Form the Lagrangian L

$$L(\underline{x}; \lambda) = f(\underset{\sim}{x}) + \sum_{i=1}^{\ell} \lambda_i h_i(\underline{x})$$

$$x \in R^n, \quad n \geq \ell$$

2. Find all $(\bar{x}, \bar{\lambda})$ satisfying the following system of nonlinear algebraic equations

Nec. Condition

$$\left\{\begin{array}{l} \underset{\sim}{\nabla} L(\underset{\sim}{x}, \underset{\sim}{\lambda}) = \nabla f(\underset{\sim}{x}) + \sum_{i=1}^{l} \lambda_i \nabla h_i(\underset{\sim}{x}) = 0 \\ \qquad (\rightarrow \text{gives } n \text{ equations}) \\ \underbrace{\left(\frac{\partial L}{\partial \lambda_i}\right)}_{\frac{\partial L}{\partial \lambda}}(\underset{\sim}{x}, \underset{\sim}{\lambda}) = h_i(\underset{\sim}{x}) = 0 \quad (i=1,\cdots,l) \\ \qquad (\rightarrow l \text{ equations}) \end{array}\right.$$

$\rightarrow$ called the Lagrangian Condition

$(\bar{x}, \bar{\lambda})$; Lagrangian points

3. Examine each solution $(\bar{x}, \bar{\lambda})$ to see if it is a minimum point (such as checking sufficient conditions)

Remark : By Lagrange's Method, constrained problems are treated as unconstrained problems.

<u>Example A</u>

$$\min \ f(\underset{\sim}{x}) = \left(x_1 - \frac{13}{3}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 - x_3$$

$$\text{s.t} \quad h_1(\underset{\sim}{x}) = x_1 + \frac{5}{3} x_2 - 10 = 0$$

$$h_2(\underset{\sim}{x}) = (x_2 - 2)^2 + x_3 - 4 = 0$$

Sol :

i) Form $L(\underset{\sim}{x}, \lambda)$

$$L(\underset{\sim}{x}, \lambda) = f(\underset{\sim}{x}) + \lambda_1 h_1(\underset{\sim}{x}) + \lambda_2 h_2(\underset{\sim}{x})$$

ii) Necessary condition ($\nabla L = 0$, $\frac{\partial L}{\partial \underline{\lambda}} = 0$)

$$\begin{cases} \frac{\partial L}{\partial x_1} = 2(x_1 - \frac{13}{3}) + \lambda_1 = 0 \\ \frac{\partial L}{\partial x_2} = 2(x_1 - \frac{1}{2}) + \frac{5}{3}\lambda_1 + 2\lambda_2 (x_2 - 2) = 0 \\ \frac{\partial L}{\partial x_3} = -1 + \lambda_2 = 0 \end{cases}$$

$$\begin{cases} \frac{\partial L}{\partial \lambda_1} = x_1 + \frac{5}{3}x_2 - 10 = 0 \\ \frac{\partial L}{\partial \lambda_2} = (x_2 - 2)^2 + x_3 - 4 = 0 \end{cases}$$

$\exists$ only one solution $\bar{x} = \{35/6,\ 5/2,\ 15/4\}^T$

$\bar{\lambda} = (-3, 1)^T$

※ In this particular case, the solution is unique but multiple solutions exist in general cases.

iii) Check if $\bar{\underline{x}}$ is a min point

1) inspect it by using a figure

2) use the sufficient condition

$H_L(\bar{\underline{x}})$ is PD = ? → will study

→ $\bar{\underline{x}}$ is indeed a min point.

Remarks: h_1, h_2 satisfies CQ for all $x \in \Omega$

since $\underline{\nabla} h_1 = \begin{pmatrix} 1 \\ 5/3 \\ 0 \end{pmatrix}$, $\underline{\nabla} h_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ are

indepedent.

Sufficient Conditions

(somewhat difficult)

(1) If constraint equations can be parameterized explicitly, the analysis is easy.

Back to Ex A. Min $f(x_1, x_2, x_3)$

s.t $h_1(\underset{\sim}{x}) = 0$ and $h_2(\underset{\sim}{x}) = 0$ ← previous example

Let $x_2 = \xi_2(t) = t$

$x_1 = \xi_1(t) = 10 - \frac{5}{3}t \leftarrow h_1(\underline{x}) = 0$

$x_3 = \xi_3(t) = 4 - (t-2)^2 \leftarrow h_2(\underline{x}) = 0$

Then $\tilde{f}(t) \triangleq f(\xi_1(t), \xi_2(t), \xi_3(t))$

$$= \underbrace{\frac{43}{9}t^2 - \frac{215}{9}t + \frac{1165}{36}}_{\text{now unconstrained problem!}}$$

$$\frac{d\tilde{f}}{dt} = \frac{86}{9}t - \frac{215}{9} = 0 \qquad t = \frac{5}{2}$$

P.D check → $\frac{d^2\tilde{f}}{dt^2} = \frac{86}{9} > 0 \Rightarrow \tilde{x}$ is a min pt.

Remark: It is difficult or impossible to parameterize general constraint equations

$\Rightarrow$ need a direct method

↳ next subject

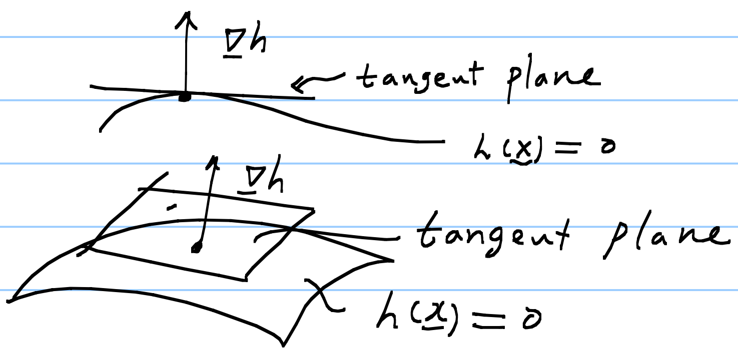
(2) Use Local Method Based on the Approximation of the Feasible set $\underset{\sim}{\Omega} = \{ \underset{\sim}{x} \mid h_i(\underset{\sim}{x}) = 0,\ i = 1, \cdots, l \}$

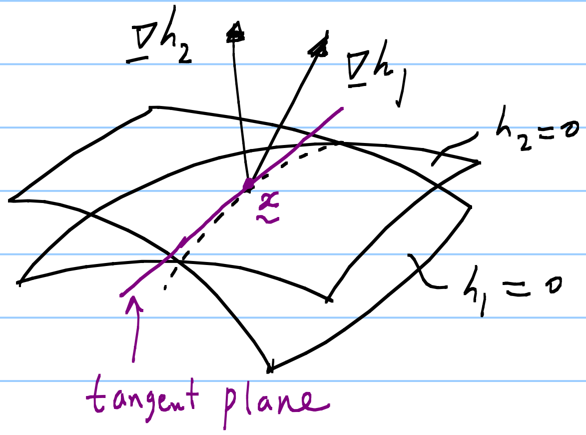
$\Rightarrow$ Approximate Ω by a tangent plane M at $\underset{\sim}{x}$

(Because we only need to check the local convexity of f near $\underset{\sim}{x}$)

$$M = \{ \underset{\sim}{d} \mid \underset{\sim}{\nabla} h^T \underline{d} = 0,\ i = 1, \cdots, l \}$$

$\underset{\sim}{d} \Rightarrow$ feasible-enough direction





CQ holds here. If not, M cannot be defined.

Sufficient Condition

If (x^*, λ^*) satisfies the Lagrange condition (FONC) and $\underline{d^T H_L(x^*) d > 0 \text{ for } d \in M, d \neq 0}$

where H_L = Hessian of $L(x, \lambda)$,

then x^* is a local min.

Back to Example A

1) find d satisfying $\nabla h_1^T d = 0$ and $\nabla h_2^T d = 0$

$$\nabla h_1 = \{1,\ 5/3,\ 0\}^T, \quad \nabla h_2 = \{0,\ 1,\ 1\}^T$$

Let $d = \{d_1,\ d_2,\ d_3\}^T$

$$\nabla h_1^T d = 0 \rightarrow d_1 + \tfrac{5}{3} d_2 = 0$$
$$\nabla h_2^T d = 0 \rightarrow d_2 + d_3 = 0$$
} one indep variable

Then one can set;

$$d_2 = t, \rightarrow d_1 = -\tfrac{5}{3}t, \quad d_3 = -t$$

$$\therefore M = \left\{ \underset{\sim}{d} \,\middle|\, \left(-\tfrac{5}{3}t,\ t,\ -t\right) \right\}^T \quad t \in \mathbb{R}$$

2) Compute H_L at $(\underline{x}^*, \underline{\lambda}^*)$ and check PD of $\underline{d}^T \underline{H}\, \underline{d}$ with $\underset{\sim}{d} \in M$

$$L = \left(x_1 - \tfrac{13}{3}\right)^2 + \left(x_2 - \tfrac{1}{2}\right)^2 - x_3$$
$$+ \lambda_1 \left(x_1 - \tfrac{5}{3}x_2 - 10\right)$$
$$+ \lambda_2 \left((x_2 - 2)^2 + x_3 - 4\right)$$

$$\underset{\sim}{H}_L = \underline{\nabla}(\underline{\nabla} L)$$
$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2+2\lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\underset{\sim}{x}^* = \left(\tfrac{35}{6}, \tfrac{5}{2}, \tfrac{15}{4}\right)^T$$
$$\lambda^* = (-3, 1)^T$$

$$\underline{d}^T \underset{\sim}{H}_L \underline{d} = 2\left(-\tfrac{5}{3}t\right)^2 + 4t^2 + 0(-t^2)$$
$$= \left(\tfrac{50}{9} + 4\right)t^2 \rightarrow \text{P-D}$$

$\therefore \quad (\underline{x}^*, \underline{\lambda}^*) \rightarrow$ Local Min

** **Remarks** : $H_L(\underline{x}^*, \underline{\lambda}^*)$ itself is not positive-definite, but $d^T \underset{\sim}{H}(\underline{x}^*, \lambda^*)\underline{d}$ for $d \in \underline{M}$, $d \neq 0$ is positive-definite for Local min.

Example B $\quad \min f(\underline{x}) = 2x_1 + x_2$

s.t. $h(\underset{\sim}{x}) = x_1^2 + x_2^2 - 1 = 0$

Lagrange point, Local min = ?

Sol : $L(\underline{x}; \underline{\lambda}) = f(\underline{x}) + \lambda\, h(\underline{x})$

$= 2x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 1)$

$\frac{\partial L}{\partial x_1} = 2 + 2\lambda x_1 = 0 \quad \rightarrow x_1 = -\frac{1}{\lambda} \quad (a)$

$\frac{\partial L}{\partial x_2} = 1 + 2\lambda x_2 = 0 \quad \rightarrow x_2 = -\frac{1}{2\lambda} \quad (b)$

$\frac{\partial L}{\partial \lambda} = x_1^2 + x_2^2 - 1 = 0 \quad (c)$

$(a, b, c) \rightarrow \quad \lambda^2 = \frac{5}{4} \quad \rightarrow \lambda = \pm\frac{\sqrt{5}}{2}$

i) $\lambda = \lambda^{(1)} = \frac{\sqrt{5}}{2}$; $\quad x_1^{(1)} = -\frac{2}{\sqrt{5}}, \quad x_2^{(1)} = -\frac{1}{\sqrt{5}}$

ii) $\lambda = \lambda^{(2)} = -\frac{\sqrt{5}}{2}$; $\quad x_1^{(2)} = +\frac{2}{\sqrt{5}}, \quad x_2^{(2)} = \frac{1}{\sqrt{5}}$

Check the sufficiency

$$H_L = \underline{\nabla}\, \underline{\nabla} f + \lambda\, \underline{\nabla}\, \underline{\nabla} h$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{bmatrix}$$

For $\lambda = \lambda^{(1)} = \sqrt{5}/2$

Compute $\underline{d}^{(1)}$ first :

$$\underline{\nabla} h^{(1)} = \begin{Bmatrix} 2x_1 \\ 2x_2 \end{Bmatrix} = \begin{Bmatrix} -4/\sqrt{5} \\ -2/\sqrt{5} \end{Bmatrix}^T$$

$\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$

$$(\underline{\nabla} h^{(1)})^T d = 0 \qquad \underline{d^{(1)} = \{1, -2\}}$$

$$\text{Check} \quad \underset{\sim}{d}^{(1)T}\, \underset{\sim}{H}^{(1)}\, \underline{d}^{(1)} = [1, -2] \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$= 5\sqrt{5} > 0$$

$\Rightarrow (\underline{x}^{(1)}, \underline{\lambda}^{(1)})$: strictly local min

For $\underset{\sim}{\lambda} = \underset{\sim}{\lambda}^{(2)} = -\sqrt{5}/2$

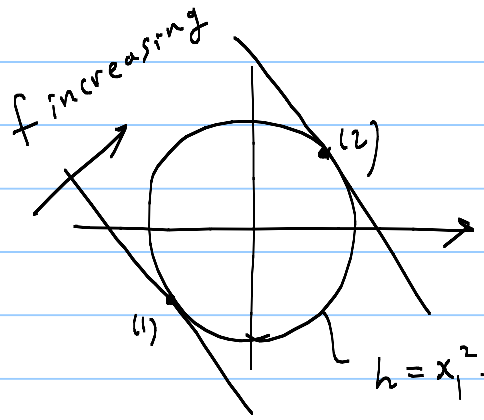
repeat the analysis done for $\underset{\sim}{\lambda} = \underset{\sim}{\lambda}^{(1)}$

$$\underset{\sim}{d} = \{1, -2\}^T$$

$$\underset{\sim}{d}^T \underset{\sim}{H}_L \underset{\sim}{d} = -5\sqrt{5} < 0, \quad \text{not a Local min}$$

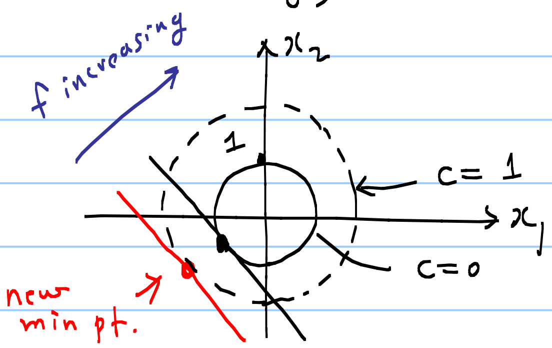
(actually Local max)

graphical Interpretation



$f = 2x_1 + x_2$

Question: What if $h(x) = 0$ changes to $h(x) = c$, say, $h(x) = 1$ (in the above example)



fmin changes as c changes!
(In this case, f decreases)

Claim: Let $\underbrace{x^* = x^*(c)}$ → means that min of $f(\underline{x})$ changes as c changes.

Then

$$\boxed{\underbrace{\left.\frac{\partial f(x^*(c))}{\partial c}\right|_{c=0}} = -\lambda^* = \left(-\frac{\sqrt{5}}{2}\right)}$$

sensitivity

"The rate of change in the function value per unit change in the constraint requirement

Example B

* In Example B, $f_{min} = f(x^*)$ will decrease further because $\partial f(\underline{x}^*(c))/\partial c = -\lambda^* < 0$.

Sensitivity Analysis

↳ how much the objective function value changes at $\underline{x}^*$ as the constraint requirements change?

Original min problem

$$\min f(\underline{x})$$
$$\text{s.t} \quad h_i(\underline{x}) = 0 \quad (i = 1, \cdots, l)$$

If $(\underline{x}^*, \lambda^*)$ minimizes $f(\underline{x}^*)$, it satisfies

$$\nabla f(\underline{x}^*) + \sum_{i=1}^{l} \lambda_i^* \underline{\nabla} h_i(x^*) = 0 \quad (a)$$
$$h_i(x^*) = 0 \quad (i = 1, \cdots, l) \quad (b)$$

Perturbed Min Problem

$$\min f(\underline{x}) = 0$$
$$\text{S.t} \quad h_i(\underline{x}) = C_i \quad (i = 1, \cdots, l)$$

If $(\underline{x}(\underline{c}), \underline{\lambda}(\underline{c}))$ is assumed to minimize the problem,

① $\underline{x}(\underline{0}) = \underline{x}^*, \quad \underline{\lambda}(0) = \underline{\lambda}$ (c)

② $$\left\{ \begin{array}{ll} \underline{\nabla} f(\underline{x}(\underline{c})) + \sum_{i=1}^{l} \lambda_i(\underline{c}) \underline{\nabla} h_i(x(\underline{c})) = 0 & (d) \\ h_j(\underline{x}(\underline{c})) = C_j \quad (j = 1, \cdots, l) & (e) \end{array} \right.$$

We want to know

$$\left.\frac{\partial f(\underline{x}(\underline{c}))}{\partial c_i}\right|_{\text{at } \underset{\sim}{c}=0} = ?$$

the sensitivity of f with respect to the change in the constraint requirement

$$\left.\frac{\partial f(\underline{x}(\underline{c}))}{\partial c_i}\right|_{\underline{c}=0} \underbrace{=}_{\text{chain rule}} \left.\sum_{k=1}^{n} \overbrace{\frac{\partial f}{\partial x_k}}^{\partial f/\partial x_k|_{\underline{x}^*}} \frac{\partial x_k(\underline{c})}{\partial c_i}\right|_{\underline{c}=0}$$

$$= \left.\sum_{k=1}^{n}\left[-\underbrace{\sum_{j=1}^{\ell}\lambda_j^* \frac{\partial h_j(\underline{x}^*)}{\partial x_k}}_{\text{By Eq. (a)}}\right]\frac{\partial x_k}{\partial c_i}\right|_{\underline{c}=0} = \sum_{j=1}^{\ell} -\lambda_j^* \overbrace{\left.\sum_{k=1}^{n}\frac{\partial h_j(\underline{x}^*)}{\partial x_k}\frac{\partial x_k}{\partial c_i}\right|_{\underline{c}=0}}^{\delta_{ji} \text{ (will be shown; *)}} \quad \text{---- (f)}$$

$$= -\sum_{j=1}^{\ell}\lambda_j^* \delta_{ji} = -\lambda_i^*$$

$$\therefore \boxed{\left.\frac{\partial f(x(\underline{c}))}{\partial c_i}\right|_{\underline{c}=0} = -\lambda_i^* \iff \underline{\nabla}_c f(\underline{x}(\underline{c})) = -\underset{\sim}{\lambda}}$$

→ gives the meaning of λ

<Proof of (*)>

$$\boxed{\sum_{k=1}^{n} \frac{\partial h_j(\underline{x}^*)}{\partial x_k} \frac{\partial x_k}{\partial c_i} = \delta_{ji}}$$

- $\displaystyle \frac{\partial}{\partial c_i} \Big[\text{LHS of Eq. (e)} \Big]_{\underline{c}=0}$

$$= \frac{\partial}{\partial c_i} h_j(\underline{x}(\underline{c}))\Big|_{\underline{c}=0} = \sum_{k=1}^{n} \frac{\partial h_j}{\partial x_k}\Big|_{\underline{x}=\underline{x}^*} \frac{\partial x_k}{\partial c_i}$$

- $\displaystyle \frac{\partial}{\partial c_i} \Big[\text{RHS of Eq. (e)} \Big] = \delta_{ji}$

$\therefore$

$$\rightarrow \boxed{\sum_{k=1}^{n} \frac{\partial h_j(\underline{x}^*)}{\partial x_k} \frac{\partial x_k}{\partial c_i} = \delta_{ji} \qquad (*)}$$