노트 제목

NLP with Inequality Constraints

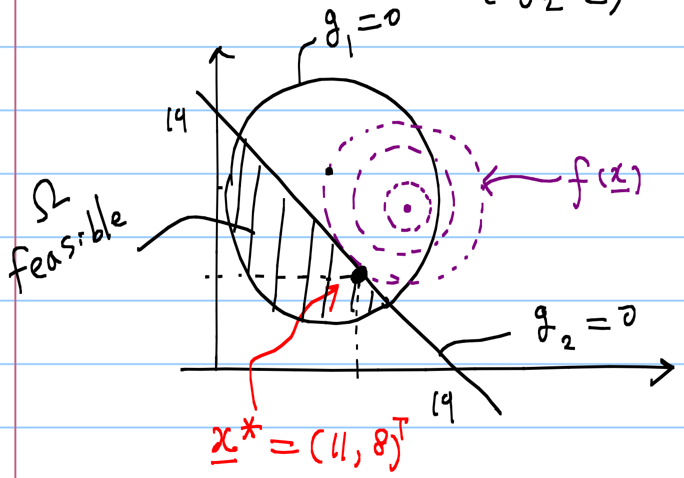
1

$$\min f(\underset{\sim}{x}), \quad x \in \mathbb{R}^n$$
$$\text{s.t} \quad g_i(\underset{\sim}{x}) \le 0 \qquad i = 1, \cdots, m$$

Example 13

(p289 E&K Fig 9.10 a)

$$\min f(\underset{\sim}{x}) = (x_1 - 14)^2 + (x_2 - 11)^2$$
$$\text{s.t} \begin{cases} g_1(\underset{\sim}{x}) = (x_1 - 11)^2 + (x_2 - 13)^2 - 7^2 \le 0 \\ g_2(\underset{\sim}{x}) = x_1 + x_2 - 19 \le 0 \end{cases}$$



Observations at $\underset{\sim}{x} = \underset{\sim}{x}^*$

1. $g_1(\underset{\sim}{x}^*) = -24 < 0$ inactive
2. $g_2(\underset{\sim}{x}^*) = 0$ active

∴ need to consider the active constraints only

⇒ Ignore Inactive Constraints

⇒ Therefore,

may consider

$$\min f(\underset{\sim}{x}) = (x_1 - 14)^2 + (x_2 - 11)^2$$
$$\text{s.t} \quad g_2(x) = x_1 + x_2 - 19 = 0$$

↑ Active Constraint (Equality Const)

$\Rightarrow$ Apply the Lagrange Method with

$$L = f(\underline{x}) + \mu_2 g_2(\underline{x})$$

$\boxed{\text{Issue}}$ How to find which constraints are active at optimality

$\Rightarrow$ "Orthogonality or Switching Conditions" are considered

Orthogonality Conditions

Back to Example B

May write L as

$$L(\underline{x}, \mu_1, \mu_2) = f(\underline{x}) + \mu_1 g_1(\underline{x}) + \mu_2 g_2(\underline{x})$$

At optimal point $\underline{x}^*$, observe that

$$\begin{cases} g_1(x^*) = -24 \neq 0 \rightarrow \text{Inactive}, & \mu_1^* = 0 \\ g_2(x^*) = 0 \Rightarrow \text{active} & \mu_2^* = 6 \neq 0 \end{cases}$$

Thus $\begin{cases} \mu_1 \, g_1 = 0 \\ \mu_2 \, g_2 = 0 \end{cases}$ at optimal points

$$\Rightarrow \quad \mu_i \, g_i(\underset{\sim}{x}) = 0 \quad i = 1, \cdots, m$$

or $\quad \underset{\sim}{\mu}^T \underset{\sim}{g}(\underset{\sim}{x}) = 0$

Additional conditions ("Orthogonality" or "switching") to be considered

KKT Conditions for NLP with Inequality Constraints

Given $f(\underset{\sim}{x}) \quad x \in R^n$

Subject to $\quad g_i(\underset{\sim}{x}) \leq 0 \quad i = 1, \cdots, m$

If the Lagrangian L is given by

$$L(\underset{\sim}{x}, \underset{\sim}{\mu}) = f(\underset{\sim}{x}) + \sum_{i=1}^{m} \mu_i \, g_i(\underset{\sim}{x}),$$

the KKT conditions are

① $\nabla L(\underset{\sim}{x}, \underset{\sim}{\mu}) = 0 \leftarrow$ Stationarity

② $\dfrac{\partial L}{\partial \mu_i}(\underset{\sim}{x}, \underset{\sim}{\mu}) \leq 0 \quad (i = 1, \cdots, m) \leftarrow$ Feasibility

$(\Leftrightarrow \; g_i(\underset{\sim}{x}) \leq 0)$

③ $\mu_i \, g_i(\underset{\sim}{x}) = 0 \quad (i = 1, \cdots, m) \leftarrow$ switching / orthogonality

* → ④ $\mu_i \geq 0 \quad (i = 1, \cdots, m)$ non negativity

※ The point $(\bar{x}, \bar{\mu})$ satisfying the KKT condition is called a KKT point.

Q: Why do we need 4 ?

↳ Some Analysis Involved

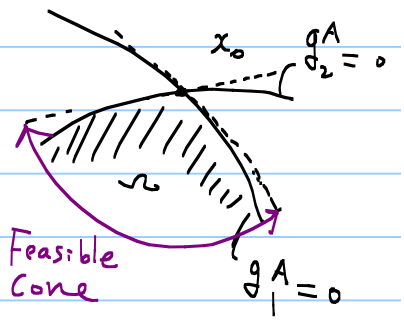
⇑ The subject of Subsequent Discussions

Some Definition

- $\underset{\sim}{d}$ is a feasible direction if

$$(-\underset{\sim}{\nabla} g_i^A)^T \underline{d} > 0 \quad (i \in I_{Active})$$

$$\text{or} \quad (\underline{\nabla} g_i)^T d < 0 \qquad (1)$$



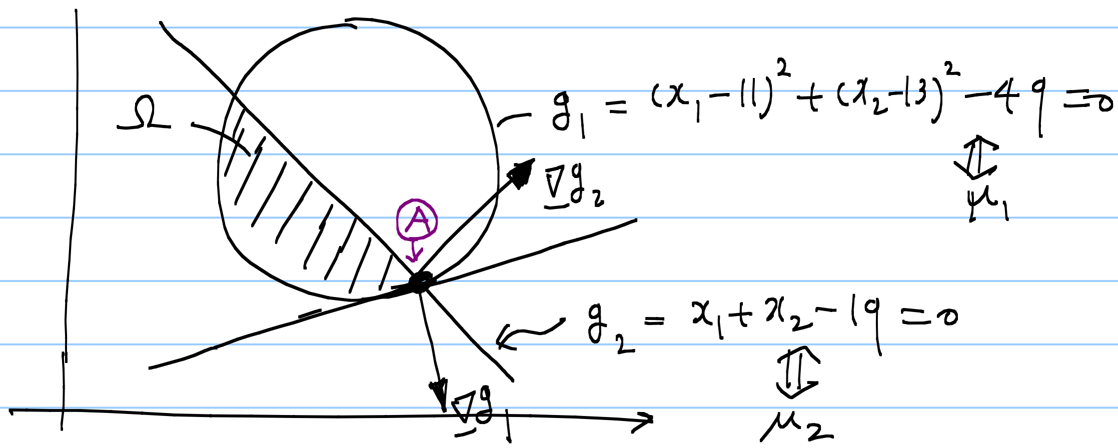
- $\underset{\sim}{d}$ is a descent direction if

$$\boxed{(-\underset{\sim}{\nabla} f)^T \underset{\sim}{d} > 0} \leftrightarrow (\nabla f)^T \underset{\sim}{d} < 0 \qquad (2)$$

** At optimal points, there is no feasible-descent direction $\Rightarrow$ I.e, there is no direction $\underset{\sim}{d}$ satisfying Conditions (1) and (2) simultaneously.

we are doing this, to explain why $\mu_i \geq 0$ is needed.

Consider the Following Situation for KKT



Assumption: $\begin{cases} g_1 \text{ and } g_2 \text{ = active at } A \\ \text{CQ holds at } A \end{cases}$

i) At A, $\quad g_1(\underline{x}_A, \underline{\mu}) \leq 0, \quad g_2(\underline{x}_A, \underline{\mu}) \leq 0 \rightarrow$ 2 okay!

$\mu_1 g_1 = 0$ $(\because g_1 = 0)$, $\quad \mu_2 g_2 = 0$ $(\because g_2 = 0)$ $\rightarrow$ 3 okay!

ii) If 1 of KKT is satisfied,

$$L = f(\underline{x}) + \mu_1 g_1(\underline{x}) + \mu_2 g_2(\underline{x})$$

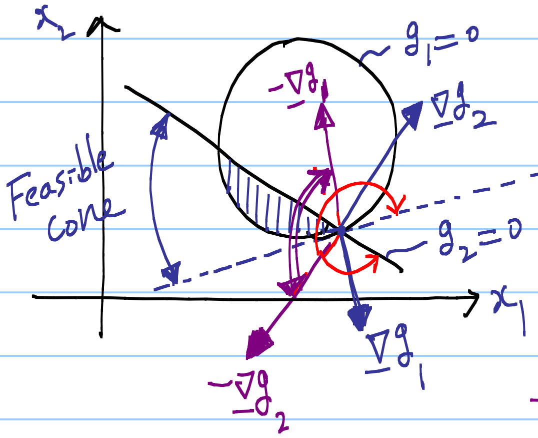
By 1: $\underline{\nabla} L = \nabla f(\underline{x}) + \mu_1 \underline{\nabla} g(x) + \mu_2 \underline{\nabla} g_2(x) = 0$

or $\quad -\underline{\nabla} f = \mu_1 \underline{\nabla} g_1 + \mu_2 \underline{\nabla} g_2 \quad$ (at A)

meaning $-\nabla f$ = linear combination of ∇g_1 and ∇g_2

⇒ There are 3 possible situations for which ①, ② and ③ are simultaneously satisfied.

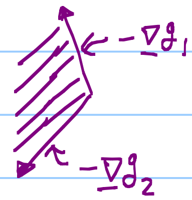
<1> Case 1 : $\mu_1, \mu_2 < 0$



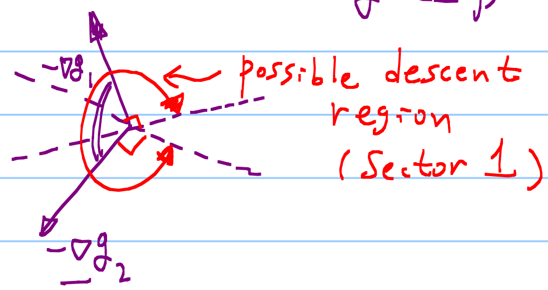
$$-\nabla f = (-\mu_1)(-\nabla g_1) + (-\mu_2)(-\nabla g_2)$$

(Linear Combination)

$\therefore$ $-\nabla f$ lies in



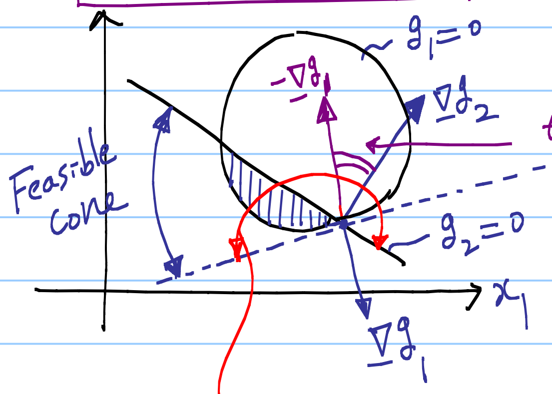
Therefore, possible descent region can be determined by $(-\nabla f)^T d > 0$



⇒ In this case, a feasible $\boxed{\text{and}}$ descent direction can be possible. ∴ Not optimal point

⟨2⟩ Case 2 ($\mu_1 < 0$ and $\mu_2 > 0$) (or $\mu_1 > 0$ and $\mu_2 < 0$)

$\boxed{\mu_1 < 0 \ \& \ \mu_2 > 0}$

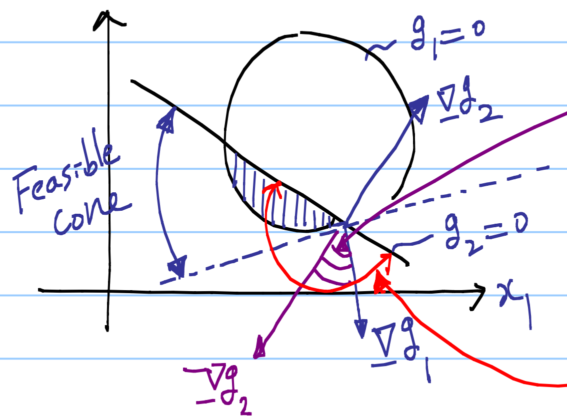


the region where
$$-\nabla f = (-\mu_1)(-\nabla g_1) + \mu_2 \nabla g_2$$
can lie.

Sector 2-1 : possible descent region $\left((-\nabla f)^T d > 0 \right)$

→ Can be a feasible $\boxed{\text{and}}$ descent direction

$\boxed{\mu_1 > 0 \ \& \ \mu_2 < 0}$

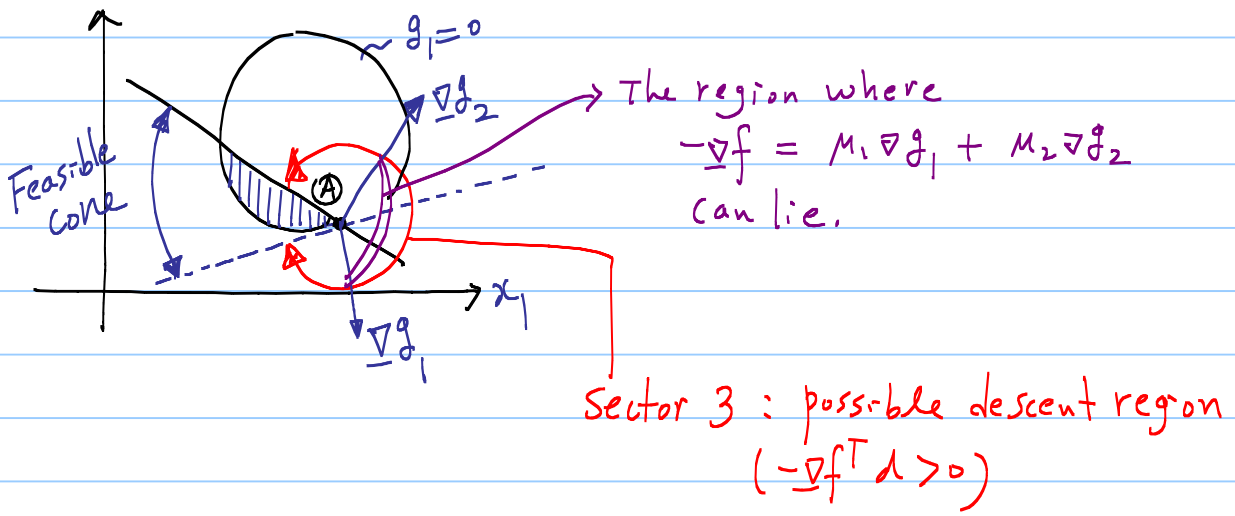


The region where
$$-\nabla f = \mu_1 \nabla g_1 + (-\mu_2)(-\nabla g_2)$$
can lie.

Sector 2-2 : possible descent direction $[(-\nabla f)^T d > 0]$

→ Can be a feasible $\boxed{\text{and}}$ descent direction

<3> Case 3 : $\mu_1 > 0$ and $\mu_2 > 0$



⇒ In this case, there can be No feasible-descent direction. ∴ Point Ⓐ is optimal

<Summary: KKT (necessary) condition>

① $\nabla L(\underline{x}, \underline{\mu}) = 0$ ← Stationarity

② $\frac{\partial L}{\partial \mu_i}(\underline{x}, \underline{\mu}) \le 0 \quad (i = 1, \cdots, m)$ ← Feasibility
$(\Leftrightarrow g_i(\underline{x}) \le 0)$

③ $\mu_i g_i(\underline{x}) = 0 \quad (i = 1, \cdots, m)$ ← switching / orthogonality

④ $\mu_i \ge 0 \quad (i = 1, \cdots, m)$ non negativity ←

Example C $\quad \min \; f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \quad$ (a)

s.t. $\quad h_1(x_1, x_2) = x_1^2 + x_2^2 - 5 \leq 0 \quad$ (b)

$\quad h_2(x_1, x_2) = 3x_1 + x_2 - 6 \leq 0 \quad$ (c)

Find KKT points.

<sol>: $\quad L(\underset{\sim}{x}; \underset{\sim}{\mu}) = f(\underset{\sim}{x}) + \mu_1 h_1(\underset{\sim}{x}) + \mu_2 h_2(\underset{\sim}{x})$

i) $\dfrac{\partial L}{\partial \underset{\sim}{x}} = 0 \qquad 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0 \quad$ (d)

$\qquad 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 = 0 \quad$ (e)

ii) $\dfrac{\partial L}{\partial \mu_i} \leq 0 \qquad h_1 \leq 0 \longleftarrow$ (b)

$\qquad h_2 \leq 0 \longleftarrow$ (c)

iii) $\underset{\sim}{\mu}^T \underset{\sim}{h} = 0 \qquad \mu_1 (x_1^2 + x_2^2 - 5) = 0 \quad$ (f)

$\qquad \mu_2 (3x_1 + x_2 - 6) = 0 \quad$ (g)

(V) $\qquad \mu_1 \geq 0$ and $\mu_2 \geq 0 \quad$ (h, i)

Solution procedures

- need to consider all possible combinations of active / passive constraints → then check the sign of the resulting Lagrange multipliers $\underset{\sim}{\mu}$ and feasibility

Let's consider the following case

$$\begin{cases} h_1 = \text{active}, & \text{then } \mu_1 \neq 0 \\ h_2 = \text{inactive}, & \text{then } \mu_2 = 0 \end{cases}$$

→ solve (d), (e) and (f)

$(\because \mu_1 \neq 0)$

$\hookrightarrow \cancel{\mu_1} (x_1^2 + x_2^2 - 5) = 0$ (1)

$\to 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 = 0$ (2)

$4x_1^2 + 2x_1 - 10 + 2\mu_1 x_1 = 0$ (3)

(1, 2, 3) : three equations for three unknowns μ_1, x_1, x_2

$$\Rightarrow \boxed{x_1 = 1, \ x_2 = 2, \ \mu_1 = 1}$$

Check ① $\mu_1 > 0$ okay, $\mu_2 = 0$

② $h_2(x_1 = 1, x_2 = 2) = 3x_1 + x_2 - 6$

$= 5 - 6 = -1 \leq 0$ okay

$\therefore (x_1, x_2) = (1, 2)$: KKT point

(sufficiency conditions will be considered later)

Example D : Solve LP problems by KKT method

$$\min f(x_1, x_2, x_3) = 6x_1 + 2x_2 + 3x_3$$

s.t.
$$x_1 + x_2 \geq 2 \rightarrow h_1 = 2 - x_1 - x_2 \leq 0$$
$$x_1 - x_2 + x_3 \geq 1 \rightarrow h_2 = 1 - x_1 + x_2 - x_3 \leq 0$$
$$x_1 \geq 0 \rightarrow h_3 = -x_1 \leq 0$$
$$x_2 \geq 0 \rightarrow h_4 = -x_2 \leq 0$$
$$x_3 \geq 0 \rightarrow h_5 = -x_3 \leq 0$$

Let
$$\begin{aligned} L(\underset{\sim}{x}, \underset{\sim}{\mu}) &= f + \mu_1 h_1 + \mu_2 h_2 + \mu_3 h_3 + \mu_4 h_4 + \mu_5 h_5 \\ &= 6x_1 + 2x_2 + 3x_3 + (2 - x_1 - x_2)\mu_1 \\ &\quad + (1 - x_1 + x_2 - x_3)\mu_2 - \mu_3 x_1 - \mu_4 x_2 \\ &\quad - \mu_5 x_3 \end{aligned}$$

<u>KKT conditions</u>

i) $\frac{\partial L}{\partial \underset{\sim}{x}} = 0$;

$$6 - \mu_1 - \mu_2 - \mu_3 = 0 \quad (1)$$
$$2 - \mu_1 + \mu_2 - \mu_4 = 0 \quad (2)$$
$$3 - \mu_2 - \mu_5 = 0 \quad (3)$$

ii) $\mu_1 \geq 0$ (4), $\mu_2 \geq 0$ (5), $\mu_3 \geq 0$ (6), $\mu_4 \geq 0$ (7), $\mu_5 \geq 0$ (8)

iii) $\mu_1 (2 - x_1 - x_2) = 0$ (9), $\mu_2 (1 - x_1 + x_2 - x_3) = 0$ (10)

$\mu_3 x_1 = 0$ (11), $\mu_4 x_2 = 0$ (12), $\mu_5 x_3 = 0$ (13)

iv) $x_1 + x_2 \geq 2$ (14), $x_1 - x_2 + x_3 \geq 1$ (15)

$x_1 \geq 0$ (16), $x_2 \geq 0$ (17), $x_3 \geq 0$ (18)

Let's try: $\begin{cases} h_1, h_2, h_5 : \text{active} \Rightarrow \mu_1 \neq 0, \mu_2 \neq 0, \mu_5 \neq 0 \\ h_3, h_4 : \text{Inactive} \Rightarrow \mu_3 = 0, \mu_4 = 0 \end{cases}$

(related to iii)

$$\left.\begin{array}{ll} h_1 \to & (9)\quad 2 - x_1 - x_2 = 0 \\ h_2 \to & (10)\quad 1 - x_1 + x_2 - x_3 = 0 \\ h_5 \to & (13)\quad x_3 = 0 \end{array}\right] \Rightarrow \left.\begin{array}{l} \begin{cases} x_1 = \frac{3}{2} \\ x_2 = \frac{1}{2} \\ x_3 = 0 \end{cases} \\ (\oplus\ \mu_3 = \mu_4 = 0) \end{array}\right\} (*)$$

Substituting (*) into ①, ②, ③

$$\left.\begin{array}{ll} ① \to & 6 - \mu_1 - \mu_2 = 0 \\ ② \to & 2 - \mu_1 + \mu_2 = 0 \\ ③ \to & 3 - \mu_2 - \mu_5 = 0 \end{array}\right] \to \begin{array}{ll} \mu_1 = 4 > 0 & \text{okay} \\ \mu_2 = 2 > 0 & \text{okay} \\ \mu_5 = 1 > 0 & \text{okay} \end{array}$$

※ If compared with the result by the Simplex Method for LP, the above solution is the optimal point.

For NLP with Equality and Inequality Constraints

Given

$$\text{Min } f(\underset{\sim}{x}) \quad x \in \mathbb{R}^n$$
$$\text{s.t} \quad g_i(\underset{\sim}{x}) \le 0, \; i = 1, \cdots, m$$
$$h_j(\underset{\sim}{x}) = 0, \; j = 1, \cdots, l$$

Optimality Condition (KKT Condition)

Define $L(\underset{\sim}{x}, \underset{\sim}{\mu}, \underset{\sim}{\lambda}) = f(\underset{\sim}{x}) + \sum_{i=1}^{m} \mu_i g_i(\underset{\sim}{x}) + \sum_{i=1}^{l} \lambda_i h_i(\underset{\sim}{x})$

KKT conditions

① $\dfrac{\partial L}{\partial \underset{\sim}{x}} = 0 \; ; \; \nabla f(\underset{\sim}{x}) + \sum_{i=1}^{m} \mu_i \nabla g_i + \sum_{i=1}^{l} \lambda_i \nabla h_i(\underset{\sim}{x}) = 0$

(stationarity)

② $\begin{cases} \dfrac{\partial L}{\partial \mu_i} = g_i \le 0 \quad (i = 1 \cdots m) \\ \dfrac{\partial L}{\partial \lambda_j} = h_j \le 0 \quad (j = 1, \cdots l) \end{cases}$ (feasibility)

③ $\mu_i g_i(\underset{\sim}{x}) = 0 \quad (i = 1, \cdots, m)$

(switching or orthogonality)

④ $\mu_i \ge 0 \quad (i = 1, \cdots, m)$

(non-negativity)

($\because$ sign of λ_i = unrestricted)

Sufficient Conditions

If $(\underline{x}^*, \underline{\lambda}^*, \underline{\mu}^*)$ satisfies the KKT condition and $\underset{\sim}{d}^T \underline{H}_L(\underline{x}^*)\, \underline{d} > 0$ for all $\underline{d} \in M$, $\underline{d} \neq 0$ then $\underline{x}^*$ is a local min where

$$M = \{ \underline{d} \mid \underline{\nabla} h_j^T \underline{d} = 0 \quad (i = 1, \cdots l) \text{ and }$$
$$\nabla \underline{g}_i^{A\,T} \underline{d} = 0 \quad (i \in I_A) \}$$

$\Rightarrow$ simple extension of the condition stated for equality-constrained optimization problems

Sensitivity Analysis

$$\min f(\underline{x})$$
$$\text{s.t}: \quad h_j(\underset{\sim}{x}) = c_j \quad (j = 1, \cdots, l)$$
$$g_i(\underline{x}) = d_i \quad (i = 1, \cdots, m)$$

Then $$\left.\frac{\partial f(x(\underline{c}, \underline{d}))}{\partial c_j}\right|_{\substack{\underline{c}=0 \\ \underline{d}=0}} = -\lambda_j^* \; ; \quad \left.\frac{\partial f(\underline{x}(\underline{c}, \underline{d}))}{\partial d_i}\right|_{\substack{\underline{c}=0 \\ \underline{d}=0}} = -\mu_i^*$$