KKT Theorem (Necessary Conditions)

Let x^* be local min. of (ICP) and assume that x^* is regular, i.e. $\nabla h_i(x^*)$, $i=1,\dots,m$, and $\nabla g_j(x^*)$, $j \in A(x^*)$, are linearly independent, Then

$\exists$ unique Lagrange multipliers $\lambda = \lambda_1, \dots, \lambda_m$ and $\mu = \mu_1, \dots, \mu_r$. s.t.:

$$\nabla_x L(x^*, \lambda, \mu) = 0$$

$$\mu_j \geq 0, \quad j = 1, \dots, r$$

$$\mu_j = 0 \quad \forall j \notin A(x^*) \equiv \mu_j g_j(x^*) = 0, \forall j$$

Furthermore, under twice diff. cont.,

$$y^T \nabla^2_{xx} L(x^*, \lambda, \mu) y \geq 0$$

$\forall y \in \mathbb{R}^n$ s.t. $\nabla h_i(x^*)^T y = 0$, $i=1,\dots,m$, $\nabla g_j(x^*)^T y = 0 \ \forall j \in A(x^*)$

<u>Sufficient Conditions</u> suppose x^*, λ, μ satisfy the first order necc. conditions above, and in addition

$$\mu_j > 0 \quad \forall j \in A(x^*).$$

and

$$y^T \nabla^2_{xx} L(x^*, \lambda, \mu) y > 0$$

$\forall y \neq 0$ s.t. $\nabla h_i(x^*)^T y = 0$ for $i = 1, \dots, m$

$$\nabla g_j(x^*)^T y = 0 \quad \forall j \in A(x^*)$$

Then x^* is a (strict) local min. of (ICP)

Example minimize $2x_1^2 + 2x_1x_2 + x_1^2 - 10x_1 - 10x_2$

$$\text{s.t.} \quad x_1^2 + x_2^2 \le 5$$
$$3x_1 + x_2 \le 6$$

$$f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$$

$$g_1(x) = x_1^2 + x_2^2 - 5, \quad g_2(x) = 3x_1 + x_2 - 6$$

$$\nabla f(x) = \begin{bmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{bmatrix}, \quad \nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

KKT First-order Necc. Conditions: $g_1(x^*) \le 0$, $g_2(x^*) \le 0$,
assuming x^* regular,

$$\nabla L(x^*, \mu) = \nabla f(x^*) + \mu_1 \nabla g_1(x^*) + \mu_2 \nabla g_2(x^*) = 0$$
$$\mu_1 \ge 0, \quad \mu_2 \ge 0$$
$$\mu_1 g_1(x^*) = 0, \quad \mu_2 g_2(x^*) = 0$$

$\nabla L(x^*, \mu) = 0$ does not reveal whether constraints are active or inactive. So need to check all possibilities.

Case 1 (1 inactive, 2 inactive) : $\mu_1 = \mu_2 = 0$.
(No need to check for regularity).

$$\nabla f(x) = 0 \Rightarrow \quad 4x_1 + 2x_2 - 10 = 0$$
$$2x_1 + 2x_2 - 10 = 0$$

Drop * on x for convenience

$$\Rightarrow \quad 2x_1 + x_2 = 5$$
$$x_1 + x_2 = 5$$

$$\Rightarrow \quad x_1 = 0, \quad x_2 = 5$$

But $x_1^2 + x_2^2 = 25 \not\le 5 \Rightarrow$ not feasible

Case 2 (1 inactive, 2 active), i.e. $\mu_1 = 0$, $g_2(x) = 0$

$$\nabla g_2(x) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \Rightarrow \text{all feasible } x \text{ are regular}$$

$$\nabla L(x,\mu) = \nabla f(x) + \mu_2 \nabla g_2(x) = 0$$

$$\left.\begin{aligned} \Rightarrow \quad & 4x_1 + 2x_2 - 10 + 3\mu_2 = 0 \\ & 2x_1 + 2x_2 - 10 + \mu_2 = 0 \\ g_2(x) = 0 \Rightarrow \quad & 3x_1 + x_2 - 6 = 0 \end{aligned}\right\}$$

Solving we get.

$x_1 = 2/5$

$x_2 = 24/5$

$\mu_2 = -2/5$

But $\mu_2 < 0$ not allowed $\Rightarrow$ solution invalid

Case 3 (1 active, 2 inactive), i.e., $\mu_2 = 0$, $g_1(x) = 0$

$$\nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \text{all feasible } x \neq 0 \text{ are regular.}$$

$$\nabla L(x,\mu) = \nabla f(x) + \mu_1 \nabla g_1(x) = 0$$

$$\Rightarrow \quad 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0$$

$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 = 0$$

$$g_1(x) = 0 \Rightarrow \quad x_1^2 + x_2^2 - 5 = 0$$

Solving these 3 equations on Matlab yields

$x_1^* = 1$, $x_2^* = 2$, $\mu_1 = 1$

as only real solution

Check: $x^* = (1,2)$ is regular (since $\neq 0$)

$$g_1(x^*) = 1 + 4 - 5 = 0$$

$$g_2(x^*) = 3 + 2 - 6 = -1 < 0$$

$x^* = (1,2)$, $\mu = (1,0)$ satisfy 1st order KKT.

Case 4 (1 active, 2 active) $g_1(x)=0$, $g_2(x)=0$

$$\nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \text{ and } \nabla g_2(x) = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ are l.i.,}$$

as long as $x \neq 0$, $x_1 \neq 3x_2$.

But $x=0$ does not satisfy $g_1(x)=0$

and $x_1 = 3x_2$, $g_2(x)=0 \Rightarrow 10x_2 = 6 \Rightarrow x_2 = \frac{3}{5}, x_1 = \frac{9}{5}$

and $\left(\frac{9}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = \frac{90}{25} < 5$ does not satisfy $g_1(x)=0$

$\Rightarrow$ all feasible x in this case are regular

$$\nabla L(x,\mu) = \nabla f(x) + \mu_1 \nabla g_1(x) + \mu_2 \nabla g_2(x) = 0$$

$$\Rightarrow \begin{aligned} 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 &= 0 \\ 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 &= 0 \end{aligned}$$

$g_1(x)=0 \Rightarrow x_1^2 + x_2^2 - 5 = 0$

$g_2(x)=0 \Rightarrow 3x_1 + x_2 - 6 = 0$

Solving these 4 equations on Matlab yields:

$x_1^* = 2.2$, $x_2^* = -0.5$, $\mu_1 = -2.4$, $\mu_2 = 4.2$ — (1)

or $x_1^* = 1.4$, $x_2^* = 1.7$, $\mu_1 = 1.4$, $\mu_2 = -1.0$ — (2)

(1) not valid since $\mu_1 < 0$, (2) not valid since $\mu_2 < 0$

Thus only candidate for local min is from Case 3:

$$x^* = (1,2), \quad \mu = (1,0).$$

Is $x^* = (1,2)$ a local min? check sufficiency

$$f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$$

$$g_1(x) = x_1^2 + x_2^2 - 5, \quad g_2(x) = 3x_1 + x_2 - 6$$

$$\nabla f(x) = \begin{bmatrix} 4x_1 + 2x_2 - 10 \\ 2x_1 + 2x_2 - 10 \end{bmatrix}, \quad \nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

For $x^* = (1, 2)$, $\mu = (1, 0)$,

1) $\nabla^2 L(x^*, \mu) = \nabla^2 f(x^*) + 1 \cdot \nabla^2 g_1(x^*) + 0$

$$= \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} > 0$$

2) active constraint has $\mu_1 > 0$

$\Rightarrow$ x^* is a strict local min. by suff. condition

Is x^* a global min for ICP?

Constraint set: $\mathcal{C} = \{x : g_1(x) \le 0, \ g_2(x) \le 0\}$

$$= \{x : g_1(x) \le 0\} \cap \{x : g_2(x) \le 0\}$$

$$= \underbrace{\{x : x_1^2 + x_2^2 \le 5\}}_{\text{compact}} \cap \underbrace{\{x : 3x_1 + x_2 \le 6\}}_{\text{closed}}$$

$\Rightarrow$ $\mathcal{C}$ is compact

$\Rightarrow$ By WT global min exists and $= x^*$.

General Sufficiency Condition

$$\text{minimize } f(x)$$

$$\text{s.t.} \quad x \in \mathcal{X} \leftarrow \text{possible additional constraints.}$$

$$h_i(x) = 0, \quad i = 1, \dots, m \qquad \text{--- (P)}$$

$$g_j(x) \le 0, \quad j = 1, \dots, r$$

Theorem Let

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{r} \mu_j g_j(x)$$

Suppose (x^*, λ, μ) satisfy:

$$h_i(x^*) = 0, \quad i = 1, \dots, m$$

$$g_j(x^*) \le 0, \quad j = 1, \dots, r$$

$$\mu_j \ge 0, \quad j = 1, \dots, r$$

$$\mu_j g_j(x^*) = 0, \quad j = 1, \dots, r$$

and

$$L(x^*, \lambda, \mu) = \min_{x \in \mathcal{X}} L(x, \lambda, \mu)$$

Then, x^* is a global min for (P) (linear + const)

If $\mathcal{X}$ is a convex set, and f is convex, h_i's are affine, g_j's are convex over $\mathcal{X}$, then $L(x, \lambda, \mu)$ is convex $\Rightarrow$ $\nabla L(x^*, \lambda, \mu) = 0$ is sufficient for x^* to be global min for (P).

Application of General Sufficiency Condition to example:

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad 2x_1^2 + 2x_1x_2 + x_1^2 - 10x_1 - 10x_2$$

$$\text{s.t.} \quad x_1^2 + x_2^2 \leq 5 \qquad \text{— } (P)$$

$$3x_1 + x_2 \leq 6$$

$\uparrow \mathcal{S}$

$$f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$$

$$g_1(x) = x_1^2 + x_2^2 - 5, \quad g_2(x) = 3x_1 + x_2 - 6$$

$$L(x, \mu) = f(x) + \mu_1 g_1(x) + \mu_2 g_2(x)$$

For $x^* = (1, 2)$, $\mu = (1, 0)$,

$L(x, \mu) = f(x) + g_1(x)$ is convex in x

and $\nabla L(x^*, \mu) = 0$

$$\Rightarrow \quad L(x^*, \mu) = \min_{x \in \mathbb{R}^2} L(x, \mu)$$

Thus, by general suff. condition,

x^* is global min for (P)

Example maximize $y^T x$ $\quad Q \succ 0 \quad -(P)$

s.t. $x^T Q x \leq 1$ $\quad y \neq 0$

Let $f(x) = -y^T x$, $g(x) = x^T Q x - 1$

$(P) \equiv$ minimize $f(x)$ s.t. $g(x) \leq 0$.

$\nabla f(x) = -y$ and $\nabla g(x) = 2Qx$

$\Rightarrow$ all (feasible) $x \neq 0$ are regular.

First-order KKT: $L(x,\mu) = f(x) + \mu g(x)$

$$\underbrace{\nabla_x L(x^*, \mu) = 0}, \quad \mu \geq 0, \quad \mu g(x^*) = 0$$

$$\hookrightarrow \Rightarrow -y + 2\mu Q x^* = 0 \Rightarrow y = 2\mu Q x^* \quad -(1)$$

$$\mu g(x^*) = 0 \Rightarrow \mu(x^{*T} Q x^* - 1) = 0 \quad -(2)$$

$y \neq 0 \Rightarrow$ by (1), $x^* \neq 0$, $\mu \neq 0$

$\Rightarrow$ by (2), $x^{*T} Q x^* = 1$

$\Rightarrow y^T x^* = 2\mu x^{*T} Q x^* = 2\mu$

$$\Rightarrow \quad x^* = \frac{Q^{-1} y}{2\mu}$$

$$\mu = \frac{y^T x^*}{2} = \frac{y^T Q^{-1} y}{4\mu}$$

$\Rightarrow 4\mu^2 = y^T Q^{-1} y$ and $\mu = \pm \dfrac{\sqrt{y^T Q^{-1} y}}{2}$

Since we need $\mu > 0$, $\mu = \frac{1}{2}\sqrt{y^T Q^{-1} y}$

$$x^* = \frac{Q^{-1}y}{\sqrt{y^TQ^{-1}y}}, \quad \mu = \frac{1}{2}\sqrt{y^TQ^{-1}y} \quad \text{unique}$$

Candidate for local min.

Sufficiency $\quad L(x,\mu) = -y^Tx + \mu(x^TQx - 1)$

$$\nabla L(x,\mu) = -y + 2\mu Qx$$

$$\nabla^2 L(x^*,\mu) = 2\mu Q \succ 0$$

$\Rightarrow$ x^* is unique local min.

Constraint set is compact $\Rightarrow$ x^* is global min by WT.

Alternatively, $L(x,\mu)$ is convex in x, and

$\nabla L(x^*,\mu) = 0$

$$\Rightarrow L(x^*,\mu) = \min_{x \in \mathbb{R}^n} L(x,\mu)$$

$\Rightarrow$ x^* is global min by general suff. condition

Maximum value of y^Tx

$$y^Tx^* = \frac{y^TQ^{-1}y}{\sqrt{y^TQ^{-1}y}} = \sqrt{y^TQ^{-1}y}$$