Lagrange function

Let us consider

$$(P) \quad \min_{x \in C} f_0(x) \text{ subject to } f_1(x) \leq 0$$

We may look at a family $(P_y)_{y \geq 0}$ of auxiliary problems:

$$(P_y) \quad \min_{x \in C} \left[f_0(x) + y f_1(x) \right]$$

where y plays the role of "weighting" the constraint:

Let $x^*(y)$ be optimal solution of (P_y).

Extreme cases:

- $x^*(0)$ may violate $f_1(x^*(0)) \leq 0$ unless $f_1(x) \leq 0$ was unnecessary.
- y sufficiently large guarantees $f_1(x^*(0)) \leq 0$ but $x^*(y)$ may be far from optimal

$\to$ It is desired to find an intermediate value $\bar{y}$ s.t. $\bar{y} f_1(x^*(\bar{y})) = 0$ so that $x^*(\bar{y})$ is also a solution of (P).

DEF: i) Let $D := \{ y \in \mathbb{R}^m \mid y_i \geq 0 \; i = 1..p \}$

We call $L: C \times D \to \mathbb{R}$ Lagrange function of (P):

$$L(x, y) := f_0(x) + \sum_{i=1}^{m} y_i f_i(x)$$

(ii) We call $(\bar{x}, \bar{y}) \in C \times D$ saddle point of L if $\forall x \in C, y \in D$:

$$L(x, \bar{y}) \geq L(\bar{x}, \bar{y}) \geq L(\bar{x}, y)$$

THM: Let (P) fulfil (Aconvex). Then:

i) $(\bar{x},\bar{y}) \in C \times D$ saddle point of L

$\Rightarrow$ $\bar{x}$ optimal for (P) and

$\bar{y}_i f_i(\bar{x}) = 0$, i.e., $L(\bar{x},\bar{y}) = f_0(\bar{x})$

ii) If $\bar{x}$ is optimal for (P) $\wedge$ (Aslater)

$\Rightarrow$ $\exists\, \bar{y} \in D$ such that $(\bar{x},\bar{y})$ saddle point of L

iii) If $\alpha := \inf\limits_{x \in S} f_0(x) \in \mathbb{R}$ $\wedge$ (Aslater)

$\Rightarrow$ $\exists\, \bar{y} \in D$ such that

$$\alpha = \inf_{x \in C} L(x,\bar{y}) = \max_{y \in D} \inf_{x \in C} L(x,y)$$

REM Under differentiability

$(\bar{x},\bar{y})$ saddle point of L

$\Leftrightarrow$ KKT conditions hold for $(\bar{x},\bar{y})$

For this note that $L(x,\bar{y}) \geq L(\bar{x},\bar{y})$ implies local min. of L, and hence, KKT (I).

Proof: i) $(\bar{x},\bar{y})$ saddle point of L

$\Leftrightarrow$ $\forall x \in C, y \in D$:

$$L(x,\bar{y}) \overset{(1)}{\geq} L(\bar{x},\bar{y}) \overset{(2)}{\geq} L(\bar{x},y)$$

$$\rightarrow f_0(\bar{x}) + \sum_{i=1}^{p} y_i f_i(\bar{x}) + \sum_{j=p+1}^{m} y_j f_j(\bar{x})$$

From $\bar{x} \in C$, $y \in D$ and (2) it follows

for $i = 1 \ldots p$:

$y_i \to +\infty \Rightarrow f_i(\bar{x}) \leq 0$

for $j = p+1 \dots m$

$$\left.\begin{array}{l} y_j \to +\infty : f_j(\bar{x}) \le 0 \\ y_j \to -\infty : f_j(\bar{x}) \ge 0 \end{array}\right\} \Rightarrow f_j(\bar{x}) = 0$$

$$\cancel{f_0(\bar{x})} + \sum_{i=1}^{p} \bar{y}_i f_i(\bar{x}) + \cancel{\sum_{j=p+1}^{m} \bar{y}_j f_j(\bar{x})}^{0}$$

$$\overset{(2)}{\ge} \cancel{f_0(\bar{x})} + \cancel{\sum_{i=1}^{p} y_i f_i(\bar{x})}^{0} + \cancel{\sum_{j=p+1}^{m} y_j f_j(\bar{x})}^{0} \quad 0$$

$$\Rightarrow \quad \bar{y}_i f_i(\bar{x}) = 0$$

From (1) we get and and $x \in S$

$$\underbrace{f_0(x) + \sum_{i=1}^{p} \bar{y}_i f_i(x) + \sum_{j=p+1}^{m} \bar{y}_j f_j(x)}_{} \ge f_0(\bar{x})$$

$f(\bar{x}) \ge$ because $f_j(x) = 0 \ \forall x \in S$, $f_i(x) \le 0$

Hence $\bar{x}$ optimal solution on S.

ii) $\bar{x}$ optimal solution of (P)

$$\Leftarrow \forall x \in S : \ f_0(\bar{x}) \le f_0(x)$$

$$f_i(\bar{x}) \le 0 \quad \forall i = 1 \dots p$$

$$f_j(\bar{x}) = 0 \quad \forall j = p+1 \dots m$$

$$(A\,\text{convex}) \wedge (A\,\text{slater}) \overset{\text{LEM}(ii)}{\Longrightarrow}$$

$\exists \bar{y} \in D$ s.t. $\forall x \in C$

$$L(x, \bar{y}) = f_0(x) + \sum_{i=1}^{p} \bar{y}_i f_i(x) + \sum_{j=p+1}^{m} \bar{y}_j f_j(x)$$

$$\ge f(\bar{x})$$

for $x = \bar{x} \Rightarrow \sum_i \bar{y}_i f_i(\bar{x}) \ge 0$ but

$$\bar{y}_i \ge 0, \ f_i(\bar{x}) \le 0 \Rightarrow \bar{y}_i f_i(\bar{x}) = 0$$

$$\Rightarrow L(x, \bar{y}) \ge f_0(\bar{x}) = L(\bar{x}, \bar{y})$$

$$\ge f_0(\bar{x}) + \sum_{i=1}^{p} y_i f_i(\bar{x}) + \sum_{i=p+1}^{m} y_j f_j(\bar{x})$$

$$= L(\bar{x}, y).$$

(iii) $\alpha := \inf_{x \in S} f(x) \in \mathbb{R}$

$$(A\text{convex}) \wedge (A\text{slater}) \xrightarrow{\text{LEM2}} \exists\, \bar{y} \in D$$

$$\forall x \in C: \quad L(x, \bar{y}) \geq \alpha$$

Hence:

- $$\inf_{x \in C} L(x, \bar{y})$$
$$= \inf_{x \in C} \left[f_0(x) + \sum_{i=1}^{p} \bar{y}_i f_i(x) + \sum_{i=p+1}^{n} \bar{y}_i f_i(x) \right]$$
$$= \alpha$$

Furthermore:

- $$\sup_{y \in D} L(x, y) = \begin{cases} f_0(x) & \text{for } x \in S \\ \infty & \text{else} \end{cases}$$

$$\Rightarrow \inf_{x \in C} \sup_{y \in D} L(x, y) = \inf_{x \in S} f(x) = \alpha$$

Therefore:

$$\alpha = \inf_{x \in C} \sup_{y \in D} L(x, y) \geq \sup_{y \in D} \inf_{x \in C} L(x, y)$$
$$\geq \inf_{x \in C} L(x, \bar{y}) = \alpha$$

$$\Rightarrow \alpha = \inf_{x \in C} L(x, \bar{y})$$
$$= \sup_{y \in D} \inf_{x \in S} L(x, y)$$
$$= \max_{y \in D} \inf_{x \in S} L(x, y)$$

$\square$

<u>Summary</u>: The take away of the above theorem for us is:

(1) Optimality can be encoded by means of saddle points of Lagrange functions

(2) For (P) fulfilling $(A_{convex}) \wedge (A_{slater})$ and $\alpha := \inf_{x \in S} f_0(x) \in \mathbb{R}$

$\Rightarrow \exists\, \bar{y} \in D$ such that

$$\alpha = \inf_{x \in C} L(x, \bar{y}) = \underbrace{\max_{y \in D} \inf_{x \in C} L(x,y)}$$

$\triangleright$ gives rise to a "dual program".

<u>Duality</u>

S, C, D as above

Primal form

$$L(x,y) = f_0(x) + \sum_{i=1}^{p} y_i f_i(x) + \sum_{i=p+1}^{n} y_i f_i(x)$$

Dual form

$$\tilde{L}(y) := \inf_{x \in C} L(x,y)$$

Say $\alpha := \inf_{x \in S} f_0(x) = f_0(x^*)$ then $x^* \in S$

$$\tilde{L}(y) = \inf_{x \in C} L(x,y) \leq L(x^*, y) \leq f(x^*) = \alpha$$

Hence, optimal values of the dual form $\delta := \sup_{y \in D} \tilde{L}(y)$ fulfil

$\delta \leq \alpha$, $\delta - \alpha$ is called duality gap which vanishes if (A_{slater}).