Supplements 5 to Chapter III

E. Gekeler
11/10/06

Duality Theory for the minimum problem with $h = 0$,

$$\{f(x)\,;\ x \in \mathcal{C}\,,\ -g(x) \in \mathcal{K}\} = \min! \tag{1}$$

(a) In the following both results let $\mathcal{A}\,,\ \mathcal{B}$ be arbitrary sets and

$$f : \mathcal{A} \times \mathcal{B} \ni (x, y) \mapsto f(x, y) \in \mathbb{R}$$

an arbitrary function.

Lemma 1

$$\sup{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y) \le \inf{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x, y).$$

Proof. See [Ekeland]. At first, we have

$$\forall\ z \in \mathcal{A},\ \forall\ y \in \mathcal{B} :\ \inf{}_{x\in\mathcal{A}} f(x, y) \le f(z, y).$$

Consequently

$$\forall\ z \in \mathcal{A} :\ \sup{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y) \le \sup{}_{y\in\mathcal{B}} f(z, y)$$

and therefore also

$$\sup{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y) \le \inf{}_{z\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(z, y). \quad \Box$$

Definition 1 *A pair $(x^*, y^*) \in \mathcal{A} \times \mathcal{B}$ is a saddlepoint of f if*

$$\forall\ x \in \mathcal{A},\ \forall\ y \in \mathcal{B} :\ f(x^*, y) \le f(x^*, y^*) \le f(x, y^*). \tag{2}$$

Lemma 2 *The function f has a saddlepoint in $\mathcal{A} \times \mathcal{B}$ if and only if*

$$\max{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y) = \min{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x, y). \tag{3}$$

Here, max instead sup says that the supremum is really attained.
Proof. [Ekeland].
(1°) Let there exist a saddlepoint (x^*, y^*) then by (2)

$$\sup{}_{y\in\mathcal{B}} f(x^*, y) = f(x^*, y^*) = \inf{}_{x\in\mathcal{A}} f(x, y^*). \tag{4}$$

But

$$\inf{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x, y) \le \sup{}_{y\in\mathcal{B}} f(x^*, y), \tag{5}$$

and

$$\inf{}_{x\in\mathcal{A}} f(x, y^*) \le \sup{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y). \tag{6}$$

Therefore also, by (4),

$$\inf{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x, y) \le \sup{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x, y). \tag{7}$$

Now, by Lemma 1, it follows that equation (7) is valid with the equality sign. This result together with (4) shows that also (5) and (6) are valid with the quality sign. By this way we obtain

$$\begin{aligned} f(x^*, y^*) &= \sup{}_{y\in\mathcal{B}} f(x^*, y) = \min{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x, y) \\ &= \inf{}_{x\in\mathcal{A}} f(x, y^*) = \max_{y\in\mathcal{B}} \inf_{x\in\mathcal{A}} f(x, y). \end{aligned}$$

(2°) Let conversely (3) be fulfilled, let the minimum be attained in x^* and the maximum in y^*. Then obviously

$$\inf{}_{x\in\mathcal{A}} f(x,y^*) \le f(x^*,y^*) \le \sup{}_{y\in\mathcal{B}} f(x^*,y),$$

and both inequalites are even equalities. Therefore (x^*,y^*) is a saddlepoint. □
Accordingly, for a saddlepoint (x^*,y^*),

$$\max{}_{y\in\mathcal{B}} \inf{}_{x\in\mathcal{A}} f(x,y) = f(x^*,y^*) = \min{}_{x\in\mathcal{A}} \sup{}_{y\in\mathcal{B}} f(x,y),$$

but not every point with this property is a saddlepoint as the simple example $f : \mathbb{R}^2 \ni (x,y) \mapsto f(x,y) = x\cdot y$ shows.
(b) Now return to the original problem

$$\{f(x),\ x\in\mathcal{C},\ g(x)\in -\mathcal{K}\} = \min!$$

with LAGRANGE function

$$L : \mathcal{X}\times\mathcal{Y}_d \ni (x,y) \mapsto L(x,y) = f(x) + y\circ g(x) \in \mathbb{R}$$

and consider the following three problems where $\mathcal{K}\subset\mathcal{Y}$ is the order cone and $\mathcal{K}_d\subset\mathcal{Y}$ the dual cone:
(1°) *The minimum problem (MP)*: Find a $x^*\in\mathcal{X}$ such that

$$x^* = \arg\min\{f(x),\ x\in\mathcal{C},\ g(x)\le 0\}. \tag{8}$$

(2°) *The primal* LAGRANGE *problem* (LP): Find a pair $(x^*,y^*)\in\mathcal{X}\times\mathcal{K}_d$ such that

$$L(x^*,y^*) = \min{}_{x\in\mathcal{C}} \sup{}_{y\in\mathcal{K}_d} L(x,y). \tag{9}$$

(3°) *The dual* LAGRANGE *problem* (DLP): Find a pair $(x^*,y^*)\in\mathcal{X}\times\mathcal{K}_d$ such that

$$L(x^*,y^*) = \max{}_{y\in\mathcal{K}_d} \inf{}_{x\in\mathcal{C}} L(x,y). \tag{10}$$

If the primal LAGRANGE problem and the dual LAGRANGE problem have both a *unique* solution and if both values of L are the same, (x^*,y^*) is a saddlepoint of the minimum problem by Lemma 2.

Lemma 3 (1°)

$$\forall\, x\in\mathcal{S} : f(x) = L(x,0) = \max{}_{y\in\mathcal{K}_d} L(x,y). \tag{11}$$

(2°) *Let $\mathcal{K}$ be closed, $x\in\mathcal{C}$, and let*

$$\exists\, y^*\in\mathcal{K}_d :\ L(x,y^*) = \max{}_{y\in\mathcal{K}_d} L(x,y), \tag{12}$$

then $x\in\mathcal{S}$ (i.e. x feasible) and $y^*\circ g(x) = 0$.
(3°)

$$\mathcal{S} = \emptyset \Longleftrightarrow \forall\, x\in\mathcal{C} :\ \sup{}_{y\in\mathcal{K}_d} L(x,y) = \infty.$$

Proof. See [Krabs], §2.1. (1°) Let $x\in\mathcal{S}$ then $g(x)\le 0$ hence

$$\forall\, y\in\mathcal{K}_d :\ y\circ g(x) \le 0,$$

which yields the assertion.
(2°) By assumption

$$\forall\, y\in\mathcal{K}_d :\ y^*\circ g(x^*) \ge y\circ g(x^*), \tag{13}$$

therefore $y^* \circ g(x^*) \geq 0$ for $y = 0$. But, on the other side, $y \circ g(x^*) \leq 0$ for all $y \in \mathcal{K}_d$. Else there would be a contradicition to (12) since with $y \in \mathcal{K}_d$ also $\alpha y \in \mathcal{K}_d$ for all $\alpha \geq 0$. Therefore $g(x^*) \leq 0$ by the cone corollary because $\mathcal{K}$ closed therefore x^* feasible. On setting $y = y^*$ we obtain together $y^* \circ g(x^*) = 0$.
(3°) Let $\mathcal{S}$ be non-empty then it follows from (1°) for $x \in \mathcal{S}$ that

$$f(x) = \sup{}_{y\in\mathcal{K}_d} L(x,y) < \infty.$$

Let $\mathcal{S}$ be empty and $x \in \mathcal{C}$ then $-g(x) \notin \mathcal{K}$. Then, by the cone corollary, there exists a $y^* \in \mathcal{K}_d$ such that $y^* \circ g(x) > 0$. Since $\alpha y^* \in \mathcal{K}_d$ for $\alpha \in \mathbb{R}$ (cone not pointed), we obtain $\sup{}_{y\in\mathcal{K}_d} L(x,y) = \infty$. □

Let in particular (x^*, y^*) be a solution of (LP) then $x^* \in \mathcal{C}$ and

$$L(x^*, y^*) = \max{}_{y\in\mathcal{K}_d} L(x^*, y)$$

($\sup{}_{y\in\mathcal{K}_d} L(x,y)$ is not necessarily attained for all $x \in \mathcal{C}$ by by assumption for x^*). Then, by Lemma 3(2°),

$$(L(x^*, y^*) =)\ f(x^*) = \min{}_{x\in\mathcal{C}} \sup{}_{y\in\mathcal{K}_d} L(x,y).$$

We say that the minimum problem (MP) and the Lagrange problem (LP) are *equivalent* if, for every solution x^* of (MP), there exists a $y^* \in \mathcal{K}_d$ such that (x^*, y^*) is solution of (LP) and, conversely, for every solution (x^*, y^*) of (LP) the first component x^* is a solution of (MP).

Lemma 4 *Let $\mathcal{K}$ be closed, then the minimum problem and the* Lagrange *problem are equivalent.*

Proof. (1°) Let x^* be a solution of (MP) and y^* arbitrary but such that $y^* \circ g(x^*) = 0$, e.g. $y^* = 0$. Then

$$\begin{aligned} f(x^*) &= \min{}_{x\in\mathcal{C}, g(x)\leq 0} f(x) \\ &= \min{}_{x\in\mathcal{C}, g(x)\leq 0} \sup{}_{y\geq 0}\{f(x) + y \circ g(x)\}, \\ &= \min{}_{x\in\mathcal{C}} \sup{}_{y\geq 0} L(x,y) \text{ (Lemma 3(c))} \\ &= L(x^*, y^*). \end{aligned} \tag{14}$$

Thus (x^*, y^*) is solution of (LP) for all y^* such that $y^* \circ g(x^*) = 0$.
(2°) Let (x^*, y^*) be a solution of (LP) then $y^* \circ g(x^*) = 0$ and x^* feasible by Lemma 3(2°). Moreover $f(x^*) = L(x^*, y^*)$, and the same conclusion as in (14) holds in the other direction. Therefore x^* is solution of (MP). □

As an inference to Lemma 2 we now obtain:

Corollary 1

$$\max{}_{y\in\mathcal{K}_d} \inf{}_{x\in\mathcal{C}} L(x,y) = \min{}_{x\in\mathcal{C}} \sup{}_{y\in\mathcal{K}_d} L(x,y)$$

holds if and only if the minimum problem (MP) has a saddlepoint.

(c) In global Lagrange theory we derive the *existence* of saddlepoints from the existence of a solution of the minimum problem under various assumptions. For instance, the following result is due to [Craven78], § 2.5:

Theorem 1 *Let the minimum problem (MP) be convex and solvable. Further, let* $\text{int}(\mathcal{K}) \neq \emptyset$ *and let the following* Karlin *condition be fulfilled:*

$$\forall\, 0 \neq y \in \mathcal{K}_d\ \ \exists\, x \in \mathcal{C}:\ y \circ g(x) < 0.$$

Then (MP) has a saddlepoint.

We now recall F. JOHN's set

$$\mathcal{A} = (g, f)(\mathcal{C}) + (\mathcal{K} \times \mathbb{R}_+) = \{(g(x) + k, f(x) + \varrho),\ x \in \mathcal{C},\ k \in \mathcal{K},\ \varrho \geq 0\}.$$

The set $\mathcal{S}$ of feasible points is non-empty if and only if there exists a $\sigma \in \mathbb{R}$ such that $(0, \sigma) \in \mathcal{A}$ and

$$\inf\{f(x),\ x \in \mathcal{S}\} = \inf\{\sigma,\ (0, \sigma) \in \mathcal{A}\}.$$

By this way the calculation of a solution of (MP) consists in the calculation of the smallest intersection point of $\mathcal{A}$ wit the $\mathbb{R}$-axis. The *dual problem* then consists of the calculation of a supporting hyperplane H in $\mathcal{Y} \times \mathbb{R}$ which contains $\mathcal{A}$ in the non-negative half-plane and on the other side has a maximum (and unique) intersection point $(0, \sigma)$ with the $\mathbb{R}$-axis. Such hyperplanes have the form

$$H(y, \sigma) := \{(u, \varrho) \in \mathcal{Y} \times \mathbb{R},\ \varrho + y \circ u = \sigma\}$$

where $(y, \sigma) \in \mathcal{Y}_d \times \mathbb{R}$. Let

$$H(y, \sigma)_+ := \{(u, \varrho) \in \mathcal{Y} \times \mathbb{R},\ \varrho + y \circ u \geq \sigma\}$$

denote the non-negative half-space of $H(y, \sigma)$ then

$$\mathcal{A} \subset H(y, \sigma)_+ \Longleftrightarrow \forall\, x \in \mathcal{C},\ \forall\, k \in \mathcal{K} : f(x) + y \circ (g(x) + k) \geq \sigma.$$

We now introduce the *dual-functional*

$$\varphi : \mathcal{Y}_d \ni y \mapsto \varphi(y) = \inf{}_{x \in \mathcal{C}}\{f(x) + y \circ g(x)\} \in \mathbb{R}$$

and the *dual problem* (DP)

$$\{\varphi(y),\ y \in \mathcal{T}\} = \max!,\ \mathcal{T} = \{y \in \mathcal{K}_d,\ \varphi(y) > -\infty\},$$

then the following result on *weak duality* is easily derived:

Theorem 2 (1°) *Let $x \in \mathcal{S}$ and $y \in \mathcal{T}$ then $\varphi(y) \leq f(x)$.*
(2°) *Let $\varphi(y^*) = f(x^*)$ for feasible arguments then x^* is solution of (MP) and y^* is solution of (DP).*

Proof. Because $x \in \mathcal{S}$ and $y \in \mathcal{K}_d$ we have

$$\varphi(y) \leq f(x) + y \circ g(x) \leq f(x)$$

The rest is clear. □
If

$$\varphi(y^*) < f(x^*),$$

holds for the resp. solutions of (MP) and (DP) then we speak of a *duality gap*. Further assumptions are necessary to avoid such gaps. For instance, duality gaps do *not* occur if the problem (MP) is linear or if the assumptions on the existence of a saddlepoint are fulfilled in convex nonlinear problems.

Theorem 3 *Let the minimum problem (MP) be convex, let $\mathcal{Y}$ be normed and $\mathcal{A}$ closed. Then*
(1°)
$$\mathcal{S} \neq \emptyset \text{ and } \inf_{x\in\mathcal{S}} f(x) > -\infty \Longleftrightarrow \mathcal{T} \neq \emptyset \text{ und } \sup_{y\in\mathcal{T}} \varphi(y) < +\infty.$$
In both cases, (MP) has a solution and, moreover,
$$-\infty < \sup_{y\in\mathcal{T}} \varphi(y) = \min_{x\in\mathcal{S}} f(x) < \infty.$$
(3°)
$$\mathcal{S} = \emptyset \text{ and } \mathcal{T} \neq \emptyset \Longrightarrow \sup_{y\in\mathcal{T}} \varphi(y) = +\infty.$$
(3°)
$$\mathcal{T} = \emptyset \text{ and } \mathcal{S} \neq \emptyset \Longrightarrow \inf_{x\in\mathcal{S}} f(x) = -\infty.$$

Proof see [Werner], Theorem 4.3.1.

Theorem 4 *Let the minimum problem (MP) be convex, $\mathcal{Y}$ normed, and let* $\mathrm{int}(\mathcal{A}) \cap \{0\} \times \mathbb{R} \neq \emptyset$. *Then $\mathcal{S}$ non-empty and:*
(1°) *If* $\inf_{x\in\mathcal{S}} f(x) > -\infty$*, the dual problem (DP) has a solution y^* and, moreover,*
$$\max_{y\in\mathcal{T}} \varphi(y) = \inf_{x\in\mathcal{S}} f(x).$$
(2°) *If the minimum problem (MP) has a solution x^*, then* $y^* \circ g(x^*) = 0$.

Proof see [Werner], Theorem 4.3.2.
For instance, condition $\mathrm{int}(\mathcal{A}) \cap \{0\} \times \mathbb{R} \neq \emptyset$ is fulfilled if the Slater condition is fulfilled, i.e. if $\mathrm{int}(K) \neq \emptyset$ and $g(x_0) < 0$ for some $x_0 \in \mathcal{C}$. Namely, then
$$\{(g(x) + k, f(x) + \varrho) \in \mathcal{Y} \times \mathbb{R},\ x \in \mathcal{C},\ k \in \mathrm{int}(\mathcal{K}),\ \varrho > 0\} \subset \mathrm{int}(\mathcal{A})$$
and thus in particular
$$(g(x_0) + (-g(x_0)), f(x_0) + \varrho) = (0, f(x_0) + \varrho) \in \mathrm{int}(\mathcal{A}) \cap \{0\} \cap \mathbb{R}$$
for all $\varrho > 0$. $\Box$
Remark. In analogy to the dual functional also a primal functional
$$\psi(x) = \sup_{y\in\mathcal{K}_d} \{f(x) + y \circ g(x)\}$$
may be introduced. Since $\psi(x) < \infty \Longleftrightarrow g(x) < \infty$ however this functional is not of interest, and the notion “primal functional” is used in an other context.