Laboratoire d'Arithmétique, de Calcul formel et d'Optimisation
ESA - CNRS 6090

On duality in nonconvex vector optimization in Banach spaces using augmented Lagrangians

Phan Quoc Khanh, Tran Hue Nuong & Michel Théra

Rapport de recherche n° 1998-02

Université de Limoges, 123 avenue Albert Thomas, 87060 Limoges Cedex
Tél. 05 55 45 73 23 - Fax. 05 55 45 73 22 - laco@unilim.fr

http://www.unilim.fr/laco/

ON DUALITY IN NONCONVEX VECTOR OPTIMIZATION IN BANACH SPACES USING AUGMENTED LAGRANGIANS *

Phan Quoc Khanh, Tran Hue Nuong † **Michel Théra**‡

November 15, 1997

Abstract

This paper shows how the use of penalty functions in terms of projections on the constraint cones, which are orthogonal in the sense of Birkhoff, permits to establish augmented Lagrangians and to define a dual problem of a given nonconvex vector optimization problem. Then the weak duality always holds. Using the quadratic growth condition together with the inf-stability or a kind of Rockafellar's stability called stability of degree two, we derive strong duality results between the properly efficient solutions of the two problems. A strict converse duality result is proved under an additional convexity assumption, which is shown to be essential.

Keywords : vector optimization, positively proprer minima, augmented Lagrangian, Birkhoff orthogonality, quadratic growth condition, inf-stability, stability of degree 2..

AMS subject classification : 90C29

---

*The research reported here was sponsored in part by the FICU-AUPELF program " *Coopération en mathématiques appliquées entre la Belgique, la France et le Viêtnam* ". The research of the first two authors was also supported by the the Vietnam National Basic Research Program in Natural Sciences

†Département de Mathématiques et d'Informatique, Université d'Hochiminh Ville, 227 rue Nguyen Van Cu, Q.1, Hochiminh Ville - Vietnam. Email : `pqkhanh@mathdep.edu.vn`

‡LACO, UPRESA 6060, Université de Limoges, 87060 Limoges Cedex, France, Email : `thera@cict.fr`

1 Introduction

Let X be a set, Y be a topological vector space and Z be a Banach space, Y and Z being ordered by closed convex cones K and M respectively. Let $F : X \to Y$ and $G : X \to Z$ be two mappings.
Consider the vector optimization problem

$$\min F(x)$$

$$\text{subject to} \qquad\qquad\qquad\qquad (\mathcal{P})$$

$$G(x) \in -M.$$

It is well known that whenever $Y = \mathbb{R}$ and X is a linear space, if problem $(\mathcal{P})$ is convex, i.e. F is convex and G is M-convex, one obtains the Lagrangian duality between (P) and its Lagrange dual. If $(\mathcal{P})$ is nonconvex, a nonzero duality gap appears. To derive duality results for nonconvex problems there are three major ways to overcome this duality gap.

- The first one, with probably the largest number of contributions consists of assuming generalized convexity conditions in order to prove duality for various duality schemes based on Lagrangians : Lagrange, Wolfe, Mond-Weir ... schemes, see e.g. [4], [5], [6], [11], [12], [15], [17], [19], [20].

- The second approach is based on defining general, perhaps abstract duality schemes, e.g. [7], [14].

- The third possibility to improve nonconvex situations is to use augmented Lagrangians. Since these Lagrangians are often combinations of ordinary Lagrangians and penalty functions, many mathematicians working in mathematical programming have contributed numerous results both from the theoretic and the computational point of view. We do not attempt to give a complete account here, but only to refer the reader to some basic literature on the subject; see for instance, [1], [2], [3], [9], [10], [13], [16], [18], [21].

Most of these contributions are developed in the finite dimensional setting, while references [9], [10] and [21] are devoted to problems in Hilbert spaces. Papers [9] and [10] use usual quadratic augmented Lagrangians to analyze algorithms for local solutions. An augmented Lagrangian containing multiplier terms built by orthogonal projections on the ordering cone M was proposed in [21] to investigate duality and algorithms. All the references on augmented Lagrangians cited above are related to for scalar optimization problems. The authors of this note are unaware of results on vector optimization using augmented Lagrangians, except the abstract ones in [7]. However many practical applications amount to considering vector optimization with a constraint space Z being a Banach spaces illustrated by the following example.

Example : Consider a cooperative differential game

$$\dot{y}(t) = \varphi(t, y(t), x(t)),\ y(t_0) = 0,$$

$$h_k(y(t_1)) = 0,\ g_i(t, y(t),\ x(t)) \leq 0,\ t \in [t_0, t_1],\ k = 1, \cdots, s,\ i = 1, \cdots \ell,$$

$$\int_{t_0}^{t_1} \psi_j(t, y(t),\ x(t))dt \to \inf,\ j = 1, \cdots, m.$$

Assume that the Cauchy problem involving the first two equations have a unique solution $y(x)(\cdot)$ for any admissible control $x(\cdot) \in L_\infty^m[t_0, t_1]$. Let $X = L_\infty^m[t_0, t_1], Y = \mathbb{R}^m,\ K = \mathbb{R}_+^m,\ Z = \mathbb{R}^s \times C^\ell[t_0, t_1]$

$$F(x) = \int_{t_0}^{t_1} \psi(t, y(t), x(t))dt,$$

$$G(x) = (h(y(x)(t_1),\ g(x)(\cdot)),$$

$$M = \Big\{ (h, g) \in Z \,|\ h = 0,\ g(x)(t) \geq 0,\ \forall\, t \in [t_0, t_1] \Big\}.$$

Then the game reduces to problem $(\mathcal{P})$.

The aim of this note is to consider duality for properly efficient solutions of Problem $(\mathcal{P})$. The main idea developed in order to achieve this goal is to build an augmented Lagrangian using shifted penalty functions in terms of a kind of projection on the cone M.

2 Augmented Lagrangian

Throughout the note assume that the cone M is *proximinal*, i.e.,

for each $z \in Z$, there exists $z^M \in M$ such that

$$||z - z^M|| = \min_{m \in M} ||z - m||. \tag{1}$$

Then we call z^M a *projection* of z on M. Recall that (1) is equivalent to saying that $z - z^M$ is *Birkhoff orthogonal* to the hyperplane supporting M at z^M. Here x is said to be Birkhoff orthogonal to y if

$$||x|| \leq ||x + \gamma y|| \text{ for each } \gamma \in \mathbb{R};$$

x is said to be orthogonal to a set $A \subset Z$ if x is orthogonal to all $a \in A$.

For convenience, set

$$z - z^M := z^{-M^*}, \; z - z^{-M} := z^{M^*}.$$

Note that if Z is a Hilbert space, then $z^M (z^{M^*}, z^{-M^*}$, respectively) is just the orthogonal projection of z on M (M^*, $-M^*$, respectively), where

$$M^* := \Big\{ z^* \in Z^* \mid \; < z^*, m > \geq \; 0 \;\; \forall m \in M \Big\},$$

and Z^* is the topological dual of Z.

For a given $z \in Z$, the point z^M is not unique. However, for our consideration we can take any projection of z on M for z^M.

Some properties of projections needed in the sequel are collected in the next lemma.

Lemma 1 *The following properties hold true :*

(i) For $z \notin M$, $z^M \in M$ is a projection of z on M if and only if $\mu \in -M^*$ of norm one exists such that

$$< \mu, z - z^M > = \; ||z - z^M||, \quad < \mu, z^M > = \; 0;$$

(ii) One can choose appropriate projections so that $(-z)^M = -z^{-M}$, $(-z)^{M^*} = -z^{-M^*}$;

(iii) $||z|| \geq \max\{\,||z^{M^*}||, ||z^{-M^*}||\,\}$ *;*

(iv) $||z^{M^*}|| = \min\left\{\,||m|| \;\mid\; m \in M + z\right\}$ *;*

(v) $||(x+y-m)^{M^*}|| \;\leq\; ||x^{M^*} + y^{M^*}|| \quad \forall\, x \in Z, \forall\, y \in Z, \forall\, m \in M.$

Proof - (i) By a theorem of Garkavi (e.g. [8], p. 76), z^M is a projection of z if and only if there exists $\mu \in Z^*$ of norm one such that, for each $m \in M$,

$$< \mu, z - z^M > = \; ||z - z^M||, \; < \mu, z^M > \geq < \mu, m > .$$

Since M is a cone, the inequality is equivalent to $\mu \in -M^*$ and $< \mu, z^M > = \; 0$.

(ii) is clear from the definitions.

(iii) Observe that

$$||z^{-M^*}|| \; = \; ||z - z^M|| \; = < \mu, z - z^M > = < \mu, z > \leq \; ||z||.$$

On the other hand,

$$||z^{M^*}|| = ||z - z^{-M}|| \leq ||z - (-m)|| \quad \forall\, m \in M.$$

Taking $m = 0$ one obtains $||z^{M^*}|| \leq ||z||$.

(iv)

$$\begin{aligned} ||z^{M^*}|| \;&=\; ||-z-(-z)^M|| \\ &=\; \min_{m' \in M} \; ||(-z - m')|| \\ &=\; \min_{m \in M+z} \; ||m||. \end{aligned}$$

(v) Since

$$x^{M^*} + y^{M^*} = x + (-x)^M + y + (-y)^M + m - m \in x + y + M - m,$$

in view of (iv) we have

$$\begin{aligned} ||(x+y-m)^{M^*}|| \;&=\; \min\left\{\,||m'||\,|\, m' \in M + x + y - m\right\} \\ &\leq\; ||x^{M^*} + y^{M^*}|| \end{aligned}$$

and the proof is complete. $\square$

In the sequel we shall always assume that the *dual quasi-interior* of K^*, i.e.

$$K^\bullet := \Big\{ \lambda \in Y^* \ \mid \ <\lambda, y> > \ 0 \quad \forall\, y \in K \backslash (-K) \Big\},$$

is nonempty. Let us remark that $K^\bullet \neq \emptyset$ whenever K^* has an interior or more generally when Y is Hausdorff and locally convex and K has a weakly compact base B, *i.e*, $K = \bigcup \Big\{ \lambda b \,|\, \lambda \geq 0, b \in B \Big\}$, where B is a convex set whose closure does not contain 0. For instance, any closed pointed convex cone in finite dimension has a compact base. To make use of scalar Lagrangians and to consider *positively proper minima* (p.p. minima, for short) of $(\mathcal{P})$ it is natural, for a given $\lambda \in K^\bullet$, to investigate together with $(\mathcal{P})$ the scalar problem $(\mathcal{P}_\lambda)$:

$$\min <\lambda, F(x)> \quad \text{subject to} \quad G(x) \in -M. \qquad (\mathcal{P}_\lambda)$$

At this point, we recall that a point $\bar{y} \in V \subset Z$ is said to be a *p.p. minimum* of V if there is $\lambda \in K^\bullet$ such that $\quad <\lambda, \bar{y}> \, \leq \, <\lambda, y> \quad$ for every $y \in V$.

In order to choose an appropriate augmented Lagrangian for $(\mathcal{P})$, it is worthwile to observe that $z \in -M$ if and only if $z^{M^*} = 0$ (by the definition of z^{M^*} and Lemma 1 (iv)). So an ordinary penalty functional for Problem $(\mathcal{P}_\lambda)$ is

$$\phi(x, \zeta) \ = <\lambda, F(x)> + \frac{1}{2}\zeta \ ||G(x)^{M^*}||^2.$$

It is well known that shifted penalty functionals can be used. In terms of projection, such a functional may be as follows :

$$\psi(x, \zeta, z) \ = <\lambda, F(x)> + \frac{1}{2}\zeta \ ||(G(x) - z)^{M^*}||^2.$$

Following the idea of the augmented Lagrangian in [21] we define the augmented Lagrangian for $(\mathcal{P})$ as follows :

$$L(x, \lambda, \zeta, z) \ = <\lambda, F(x)> + \frac{1}{2}\zeta||(G(x) - z)^{M^*} \ ||^2 - \frac{1}{2}\zeta||z||^2,$$

with $x \in X$, $\lambda \in K^*$, $(\zeta, z) \in \mathbb{R}_+ \times Z$. Note that the added term does not depend on x, so ψ and L are equivalent when minimized on x. Furthermore, if X is a linear space $L(\cdot, \lambda, \zeta, z)$ is convex whenever $(\mathcal{P})$ is convex. Recall that a mapping $H : X \to Y$ is said to be K-*convex* if

$$H((1-\alpha)x + \alpha y) \in (1-\alpha)H(x) + \alpha H(y) - K$$

whenever $x \in X$, $y \in X$ and α lies in $[0, 1]$. $(\mathcal{P})$ is called *convex* if F is K-convex and G is M-convex. (The convexity of $||(G(\cdot) - z)^{M^*}||$ follows from Lemma 1 (v)). Then, the attainable set of $(\mathcal{P})$ is K-convex. In the case where Z is a Hilbert space, L takes a form involving an explicit multiplier as follows. Let

$$f_\lambda(x, u) \ := \begin{cases} \langle \lambda, F(x) \rangle & \text{if} \quad G(x) - u \in -M, \\ +\infty & \text{otherwise.} \end{cases}$$

Then, by Lemma 1 (iv),

$$\begin{aligned} L(x, \lambda, \zeta, z) &= \ < \lambda, F(x) > + \frac{1}{2}\zeta \min_{n \in G(x)+M} ||u - z||^2 - \frac{1}{2}\zeta ||z||^2 \\ &= \inf_{u \in Z} \left\{ f_\lambda(x, u) + \frac{1}{2}\zeta ||u - z||^2 - \frac{1}{2}\zeta ||z||^2 \right\} \qquad (2) \\ &= \inf_{u \in Z} \left\{ f_\lambda(x, u) + < \eta, u > + \frac{1}{2}\zeta ||u||^2 \right\} \end{aligned}$$

with the multiplier $\eta := -\zeta z$.

Consider a perturbed problem $(\mathcal{P}_{\lambda,u})$ from $(\mathcal{P}_\lambda)$:

$$\min < \lambda, F(x) >, \quad \text{subject to } \ G(x) - u \in -M. \qquad (\mathcal{P}_{\lambda,u})$$

In the sequel we shall denote the optimal values of problems $(\mathcal{P}_\lambda)$ and $(\mathcal{P}_{\lambda u})$ by p_λ and $p_\lambda(u)$, respectively, while P will stand for the set of all p.p. minima of the closure of the attainable set of $(\mathcal{P})$. Then $p_\lambda(u) = \inf_{x \in X} \ f_\lambda(x, u)$ and, by (2) (true if Z is a Banach space),

$$\inf_{x \in X} L(x, \lambda, \zeta, z) = \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\zeta ||u - z||^2 - \frac{1}{2}\zeta ||z||^2 \right\}. \qquad (3)$$

3 Duality

We define the dual $(\mathcal{D})$ of $(\mathcal{P})$ as

$$\max_{y \in W} y, \qquad\qquad (\mathcal{D})$$

where

$$W = \Big\{ y \in Y \mid \exists \lambda \in K^{\bullet},\ \exists (\zeta, z) \in \mathbb{R}_+ \times Z,\ \forall x \in X,\ <\lambda, y> \leq\ L(x, \lambda, \zeta, z) \Big\}.$$

Observe that every (Pareto) maximum of $(\mathcal{D})$ is also a positively proper maximum of $(\mathcal{D})$.

Lemma 2

$$\sup_{(\zeta,z)\in\mathbb{R}_+\times Z} \Big\{ \zeta||(G(x)-z)^{M^*}||^2 - \zeta||z||^2 \Big\} = \begin{cases} 0 & \text{if} \quad G(x) \in -M, \\ +\infty & \quad \text{otherwise} \end{cases}$$

and, when $G(x) \in -M$, the supremum is attained at $(\zeta, 0) \quad \forall\, \zeta \in \mathbb{R}_+$.

Proof - If $G(x) \in -M$, by Lemma 1 (v) and (iii) one has

$$||(G(x)-z)^{M^*}|| - ||z|| = ||(-z)^{M^*}|| - ||-z|| \leq 0.$$

On the other hand, for $z = 0$,

$$G(x)^{M^*} = G(x) - G(x)^{-M} = 0,$$

i.e. the supremum is reached at $(\zeta, 0)$, $\forall\, \zeta \in \mathbb{R}_+$, and is equal to 0.
If $G(x) \notin -M$, one has

$$G(x)^{M^*} = G(x) + (-G(x))^M \neq 0.$$

Taking $(\zeta_n, z_n) = (n, 0)$ we obtain $\lim_{n\to\infty} \zeta_n ||G(x)^{M^*}||^2 = +\infty$, establishing the proof. □

Thanks to Lemma 2, as any duality scheme using Lagrangians does, the weak duality between $(\mathcal{P})$ and $(\mathcal{D})$ always holds as follows.

Proposition 1 *For any feasible point x of $(\mathcal{P})$ and any $y \in W$ of $(\mathcal{D})$ we have*

$$y \notin F(x) + K \backslash (-K).$$

Proof - Since $y \in W$, the exist $\lambda \in K^{\bullet}$ and $(\zeta, z) \in \mathbb{R}_+ \times Z$ such that

$$< \lambda, y > \leq \; L(x, \lambda, \zeta, z) \quad \forall \, x \in X.$$

Hence, by Lemma 2, $< \lambda, y > \leq < \lambda, F(x) >$ for all feasible x and we are done. $\square$

Corollary 1 *If, for a feasible $\bar{x}$, $F(\bar{x}) \in W$, then $F(\bar{x})$ is both a p.p. minimum of $(\mathcal{P})$ and a p.p. maximum of $(\mathcal{D})$.*

By virtue of Lemma 2, Problem $(\mathcal{P}_\lambda)$ can be rewritten as

$$\inf_{x \in X} \sup_{(\zeta,z) \in \mathbb{R}_+ \times Z} L(x, \lambda, \zeta, z). \qquad (`\mathcal{P}_\lambda)$$

Its dual (defined similarly to $(\mathcal{D})$ from $(\mathcal{P})$) is

$$\sup_{(\zeta,z) \in \mathbb{R}_+ \times Z} \inf_{x \in X} L(x, \lambda, \zeta, z). \qquad (\mathcal{D}_\lambda)$$

If the optimal value d_λ of $(\mathcal{D}_\lambda)$ satisfies $d_\lambda = -\infty$, then the strong duality between $(\mathcal{P}_\lambda)$ and $(\mathcal{D}_\lambda)$ means that $p_\lambda = -\infty$, which does not make any real sense. So it is natural to assume the following hypothesis (boundedness from below) :

$$\text{there exists } \; (\zeta, z) \in \mathbb{R}_+ \times Z \; \text{ such that } \; \inf_{x \in X} L(x, \lambda, \zeta, z) > -\infty.$$

This assumption is equivalent to saying (cf.[16]) that there exists $(\tilde{q}, \tilde{\zeta}) \in \mathbb{R} \times \mathbb{R}_+$, such that for each $u \in Z$ we have,

$$\tilde{q} - \frac{1}{2} \tilde{\zeta} \, ||u||^2 \leq p_\lambda(u). \tag{4}$$

Indeed, if (4) holds, taking $\zeta = \tilde{\zeta}$, $z = 0$ by (3) we see that

$$\inf_{x \in X} \; L(x, \lambda, \tilde{\zeta}, 0) = \inf_{u \in Z} \left\{ P_\lambda(u) + \frac{1}{2} \tilde{\zeta} ||u||^2 \right\} \geq \tilde{q}.$$

Conversely, if there exist $(\tilde{\zeta}, \tilde{z}) \in \mathbb{R}_+ \times Z$, $\tilde{q} \in \mathbb{R}$ such that

$$\tilde{q} \leq \inf_{x \in X} L(x, \lambda, \tilde{\zeta}, \tilde{z}),$$

then, again by (3),

$$\begin{aligned}
\tilde{q} &\leq \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\tilde{\zeta}\, ||u - \tilde{z}||^2 - \frac{1}{2}\tilde{\zeta}||\tilde{z}||^2 \right\} \\
&\leq \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\tilde{\zeta}(||u||^2 + ||\tilde{z}||^2) - \frac{1}{2}\tilde{\zeta}||\tilde{z}||^2 \right\} \\
&= \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\tilde{\zeta}||u||^2 \right\}.
\end{aligned}$$

If (4) holds, we say that Problem $(\mathcal{P}_\lambda)$ satisfies the *quadratic growth condition* (q.g.c.).

Observation : $d_\lambda = -\infty$ if and only if the q.g.c. is not satisfied for $(\mathcal{P}_\lambda)$.

The following theorem shows that the q.g.c. guarantees a classical outcome [16], nearby to strong duality, for our scheme $(\mathcal{P}_\lambda)$ and $(\mathcal{D}_\lambda)$.

Theorem 1 *$(\mathcal{P}_\lambda)$ satisfies the q.g.c. if and only if*

$$-\infty < d_\lambda = \liminf_{u \to 0} p_\lambda(u) \leq p_\lambda(0) \equiv p_\lambda. \tag{5}$$

Proof - We have to prove only the "only if" part. Consider an arbitrary $q \in \mathbb{R}$ such that $q < \liminf_{u \to 0} p_\lambda(u)$ and an $\varepsilon > 0$ small enough to have $p_\lambda(u) \geq q$ if $||u|| < \varepsilon$.
For $(\tilde{q}, \tilde{\zeta})$ in (4) one has, for all ζ sufficiently large,

$$q - \frac{1}{2}\,\zeta||u||^2 \leq \tilde{q} - \frac{1}{2}\,\tilde{\zeta}||u||^2 \text{ if } ||u|| \geq \varepsilon.$$

Then

$$q - \frac{1}{2}\,\zeta\, ||u||^2 \leq p_\lambda(u) \quad \forall\, u \in Z.$$

Hence

$$q \leq \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\zeta||u||^2 \right\}.$$

Since q is arbitrary, this implies

$$\liminf_{u \to 0} p_\lambda(u) \leq \inf_{u \in Z} \left\{ p_\lambda(u) + \frac{1}{2}\zeta||u||^2 \right\} \leq d_\lambda.$$

The opposite inequality always holds (without (4)) as follows :

$$\begin{aligned} d_\lambda &= \sup_{(\zeta,z)} \inf_{u\in Z} \Big\{ (p_\lambda(u) + \frac{1}{2}\zeta||u-z||^2 - \frac{1}{2}\zeta||z||^2 \Big\} \\ &\le \sup_{(\zeta,z)} \inf_{u\in Z} \Big\{ p_\lambda(u) + \frac{1}{2}\zeta||u||^2 \Big\} \\ &\le \sup_{\zeta} \liminf_{u\to 0} \Big\{ p_\lambda(u) + \frac{1}{2}\zeta||u||^2 \Big\} \\ &\le \sup_{\zeta} \liminf_{u\to 0} p_\lambda(u) = \liminf_{u\to 0} p_\lambda(u) \end{aligned}$$

establishing the proof. □

To obtain a corresponding result for $(\mathcal{P})$, similarly let D be the set of all p.p. maxima of the closure of the attainable set of $(\mathcal{D})$. We assume by assumption that :

$$P(u) \neq \emptyset \quad \text{for all } u \text{ in a neighbourhood of } 0.$$

Observe that

$$p_\lambda(u) = \inf_{y\in P(u)} < \lambda, y > .$$

Let us define an inferior limit for a mapping $A : Z \Longrightarrow Y$ as follows. First, for $B : Z \Longrightarrow \mathbb{R}$ we write

$$\liminf_{u\to 0} B(u) := \lim_{s\to 0} \inf_{||u||\le s} B(u).$$

Then, define

$$\begin{aligned} K^\bullet - \liminf_{u\to 0} A(u) \;\; &:= \;\; \Big\{ y \in \limsup_{u\to 0} A(u) \,|\, \exists \lambda \in K^\bullet, \\ & \qquad < \lambda, y > = \liminf_{u\to 0} < \lambda, A(u) > \Big\}, \end{aligned}$$

where the symbol "limsup" means the superior limit of sets in the Painlevé-Kuratowski sense :

$$\limsup_{u\to 0} A(u) = \bigcap_{\varepsilon>0} \overline{\bigcup_{||u||\le\varepsilon} A(u)}.$$

Then

$$\begin{aligned}\liminf_{u\to 0} p_\lambda(u) &= \lim_{s\to 0} \inf_{||u||\le s} \inf_{y\in P(u)} < \lambda, y > \\ &= \lim_{s\to 0} \inf_{||u||\le s} < \lambda, P(u) > \\ &= \liminf_{u\to 0} < \lambda, P(u) >\end{aligned}$$

and

$$K^\bullet - \liminf_{u\to 0} P(u) = \Big\{ y \in \limsup_{u\to 0} \ P(u) \,|\, \exists \lambda \in K^\bullet, \ < \lambda, y > = \liminf_{u\to 0} p_\lambda(u) \Big\}. \qquad (6)$$

If $y \in K^\bullet - \liminf_{u\to 0} P(u)$, then for any λ satisfying (6) (for this y) we say that y corresponds to λ.

We define a kind of relaxed compactness as follows. A family $A(u)$, $u \in U \subset Z$, of sets in Y is said to be $K^\bullet$-*compact* at $u = 0$ if

for each $\lambda \in K^\bullet$, for each sequence $\{u_n\}_{n\in\mathbb{N}}$ in U with limit 0, for each sequence $\{y_n\}_{n\in\mathbb{N}}$ in $clA(u_n)$ such that $< \lambda, y_n >$ converges, there exists $y \in Y$, and a subsequence $\{y_{n_k}\}_{k\in\mathbb{N}}$ converging to y as k goes to ∞.

Of course, if $A(\cdot)$ is a multifunction continuous at $u = 0$ and each $A(u)$ is relatively compact, then the family $\Big\{ A(u) \,|\, u \in Z \Big\}$ is $K^\bullet$-compact at $u = 0$. This condition is rather weak. However, there are non-$K^\bullet$-compact families as the example below shows.

Example : Let $Y = \mathbb{R}^2$, $K = \Big\{ (y_1, y_2) \,|\, -y_1 \le y_2 \le y_1 \Big\}$, $Z = \mathbb{R}$, $U = \Big\{\frac{1}{n} \,|\, n \ge 1\Big\} \cup \{0\}$, $A(0) = 1$ and $A\left(\frac{1}{n}\right) = \Big\{\left(\frac{1}{n}, n\right) \ |\, n \ge 1\Big\}$. Then, for $\lambda = (1, 0)$, $< \lambda, A\left(\frac{1}{n}\right) > = \frac{1}{n}$ tends to 0 but the sequence $\Big\{\left(\frac{1}{n}, n\right)\Big\}_{n\ge 1}$ has no convergent subsequence. So the family $A(u)$ is not $K^\bullet$-compact at $u = 0$.

Since, for Problem $(\mathcal{P})$, X is an arbitrary set and G is an arbitrary mapping, in order to obtain for $(\mathcal{P})$ a result corresponding to Theorem 1, it is essential to impose additionally the $K^\bullet$-compactness at $u = 0$ of the family of the attainable sets of $(\mathcal{P}_u)$, $u \in Z$. Also, seeking for duality results, we often assume the existence of solutions for the primal problem.

Theorem 2 *Assume that the family of the attainable sets of* $(\mathcal{P}_u)$, $u \in Z$, *is* $K^{\bullet}$*-compact at* $u = 0$. *Then the following assertions hold :*

(i) If $(\mathcal{P}_\lambda)$ *satisfies the q.g.c. and if for all* u *in a neighbourhood of zero, there exists* $y \in P(u)$ *corresponding to* λ, *then the set of* $y \in K^{\bullet} - \liminf_{u \to 0} P(u)$ *corresponding to* λ *is nonempty and contained in* D *;*

(ii) If the assumptions in (i) are satisfied for all $\lambda \in K^{\bullet}$, *then*

$$\phi \neq K^{\bullet} - \liminf_{u \to 0} P(u) \subset D \subset cl(Y \backslash (P + K \backslash - K).$$

Proof - (i) Suppose on the contrary that among all $y \in Y$ such that $< \lambda, y >= \liminf_{u \to 0} p_\lambda(u) = d_\lambda$ there is no point of $\limsup_{u \to 0} P(u)$. Then, for each such an y, for every sequence $\{u_n\}_{n \in \mathbf{N}}$ with $u_n \in Z$ and $\lim_{n \to \infty} u_n = 0$, there exists a neighbourhood $\mathcal{Y}_y$ of y, such that $P(u_n) \cap \mathcal{Y}_y = \phi$ for all large n.

On the other hand, there exists a sequence $\{u_n\}_{n \in \mathbf{N}}$ with limit 0 such that $\lim_{n \to \infty} p_\lambda(u_n) = d_\lambda$. By the assumption on $P(u_n)$ one can choose y_n in the closure of the attainable set of $(\mathcal{P}_{u_n})$ such that $u_n \in P(u_n)$ and $< \lambda, y_n >= p_\lambda(u_n)$. Hence, the $K^{\bullet}$-compactness yields the existence of $\bar{y} \in Y$ and a subsequence $\{y_{n_k}\}_{k \in \mathbf{N}}$ such that $\lim_{k \to \infty} y_{n_k} = \bar{y}$. Clearly $< \lambda, \bar{y} > = d_\lambda$. This contradicts the fact that all $y_{n_k} \notin \mathcal{Y}_{\bar{y}}$. Furthermore, it is obvious that the mentioned set of y is contained in D.

(ii) follows from (i) and the weak duality. □

Now let us return to relations between (p_λ) and (d_λ). Theorem 1 clearly has the following consequence.

Theorem 3 *Suppose that* $(\mathcal{P}_\lambda)$ *satisfies the q.g.c.. Then,* $d_\lambda = p_\lambda$, *i.e.*

$$\sup_{(\zeta,z)} \inf_x L(x, \lambda, \zeta, z) = \inf_x \sup_{(\zeta,z)} L(x, \lambda, \zeta, z)$$

if and only if $(\mathcal{P}_\lambda)$ *is inf-stable in the sense that*

$$\liminf_{u \to 0} p_\lambda(u) \geq p_\lambda.$$

For our vector problems $(\mathcal{P})$ and $(\mathcal{D})$, then the following result is valid.

Theorem 4 *Let all the assumptions in Theorem 2 be satisfied. Let, for each* $\lambda \in K^\bullet$*,* $(\mathcal{P}_\lambda)$ *be inf-stable. Then*

$$P \subset K^\bullet - \liminf_{u\to 0} P(u) \subset D.$$

Proof - Only the first inclusion has to be checked. Of course

$$P = P(0) \subset \limsup_{u\to 0} P(u).$$

On the other hand

$$P \subset \left\{ y \in Y \,|\, \exists \lambda \in K^\bullet,\ <\lambda, y> = \ p_\lambda \right\}.$$

Since $p_\lambda = \liminf_{u\to 0} p_\lambda(u)$ for all $\lambda \in K^\bullet$, the conclusion follows immediately. □

In order to have a strong duality relations between p.p. minima of $(\mathcal{P})$ and p.p. maxima of $(\mathcal{D})$ (not simply through their closures), using a traditional way (cf. [16], [21]), we must add some extra conditions.

Problem $(\mathcal{P}_\lambda)$ is said to be *locally stable of degree* 2 if there exist $\varepsilon > 0$ and $(\bar{\zeta}, \bar{z}) \in \mathbb{R}_+ \times Z$, such that $||u|| < \varepsilon$ implies

$$p_\lambda - \frac{1}{2}\bar{\zeta}\ ||u - \bar{z}||^2 + \frac{1}{2}\bar{\zeta}\ ||\bar{z}||^2 \le p_\lambda(u). \tag{7}$$

If (7) holds for all $u \in Z$, $(\mathcal{P}_\lambda))$ is called *globally stable of degree 2* (we write for short these two conditions as l.s.2 and g.s.2, respectively).

Remark : l.s.2 property is clearly stronger than the inf-stability (strictly stronger !). g.s.2 property is equivalent to the two conditions l.s.2 and q.g.c. together. Indeed, the g.s.2 property of course contains l.s.2. It implies also the q.g.c. since, by (7),

$$\begin{aligned} p_\lambda &\le \inf_{u\in Z}(p_\lambda(u) + \frac{1}{2}\bar{\zeta}|\ ||u - \bar{z}||^2 - \frac{1}{2}\bar{\zeta}\ ||\bar{z}||^2) \\ &= \inf_{x\in X}\ L(x, \lambda, \bar{\zeta}, \bar{z}). \end{aligned}$$

Conversely, by the q.g.c. property (4) one can choose $\zeta' \geq \bar{\zeta}$ ($\bar{\zeta}$ in l.s.2 property) and large enough such that for ε in l.s.2. property,

$$p_\lambda(u) + \frac{1}{2}\zeta'||u||^2 \geq p_\lambda \ \text{ if } \ ||u|| \geq \varepsilon.$$

On the other hand

$$\begin{aligned}
\inf_{||u||<\varepsilon} \Big(p_\lambda(u) + \frac{1}{2}\zeta'||u||^2\Big) &\geq \inf_{||u||<\varepsilon} \Big(p_\lambda(u) + \frac{1}{2}\bar{\zeta}\ ||u||^2\Big) \\
&\geq \inf_{||u||<\varepsilon} \Big(p_\lambda(u) + \frac{1}{2}\bar{\zeta}\ ||u-\bar{z}||^2 - \frac{1}{2}\bar{\zeta}\ ||\bar{z}||^2\Big) \\
&\geq p_\lambda.
\end{aligned}$$

So, for $\zeta = \zeta'$, $z = 0$ and for all $u \in Z$, (7) holds. □

Theorem 5 *(i) If $(\mathcal{P}_\lambda)$ is g.s.2 and $y \in P$ corresponds to λ, then y is a p.p. maximum of $(\mathcal{D})$.*

(ii) Conversely, if there is $y \in W$ belonging to the closure of the attaimable set of $(\mathcal{P})$, then, for any $\lambda \in K^\bullet$ corresponding to y (in the definition of W), $(\mathcal{P}_\lambda)$ is g.s.2.

Proof - (i) By (7) and (3) one has

$$\begin{aligned}
< \lambda, y >= p_\lambda &\leq \inf_{u\in Z} \Big(p_\lambda(u) + \frac{1}{2}\bar{\zeta}\ ||u-\bar{z}||^2 - \frac{1}{2}\bar{\zeta}||\bar{z}||^2\Big) \\
&= \inf_{x\in X}\ L(x,\lambda,\bar{\zeta},\bar{z}).
\end{aligned}$$

So $y \in W$ and $< \lambda, y >= d_\lambda$, then y is a p.p. maximum of $(\mathcal{D})$.

(ii) For the given y and λ and for $(\bar{\zeta}, \bar{z})$ from the definition of y in W we have (by Lemma 2)

$$\begin{aligned}
< \lambda, y > &\leq \inf_{x\in X} L(x,\lambda,\bar{\zeta},\bar{z}) \\
&\leq \inf \Big\{ < \lambda, F(x) > \ \mid\ G(x) \in -M \Big\} \\
&= p_\lambda.
\end{aligned}$$

Since y is in the mentioned closure, the inequalities must be equalities. Then, by (3), we have (7) for each $u \in Z$. $\square$

Corollary 2 *In order that the duality relation*

$$\inf_x \sup_{(\zeta,z)} L(x,\lambda,\zeta,z) = \max_{(\zeta,z)} \inf_x L(x,\lambda,\zeta,z)$$

hold it is necessary and sufficient that $(\mathcal{P}_\lambda)$ *be g.s.2.*

Theorem 6 *(i) If* $(\bar{x},\bar{\lambda},\bar{\zeta},\bar{z})$ *is a saddle point of* $L(x,\bar{\lambda},\zeta,z)$*, i.e., for each* $x \in X, (\zeta,z) \in \mathbb{R}_+ \times Z$*, we have*

$$L(\bar{x},\bar{\lambda},\zeta,z) \leq L(\bar{x},\bar{\lambda},\bar{\zeta},\bar{z}) \leq L(x,\bar{\lambda},\bar{\zeta},\bar{z}), \tag{8}$$

then $\bar{x}$ *is a p.p. minimizer of* $(\mathcal{P})$ *and all* $y \in Y$ *such that*

$$< \bar{\lambda}, y > = L(\bar{x},\bar{\lambda},\bar{\zeta},\bar{z}) \tag{9}$$

are p.p. maxima of $(\mathcal{D})$.

(ii) Conversely, if $\bar{x}$ *is a p.p. minimizer of* $(\mathcal{P})$ *corresponding to* $\bar{\lambda} \in K^\bullet$ *and* $(\mathcal{P}_{\bar{\lambda}})$ *is g.s.2, then there exists* $(\bar{\zeta},\bar{z}) \in \mathbb{R}_+ \times Z$ *such that (8) holds. (In fact (8) holds if and only if all* y *satisfying (9) are p.p. maxima of* $(\mathcal{D})$*.)*

Proof - (i) Lemma 2 together with (8) imply that $\bar{x}$ is feasible and, for all feasible points x,

$$\begin{aligned} < \bar{\lambda}, F(x) > &= \sup_{(\zeta,z)} L(x,\bar{\lambda},\zeta,z) \\ &\geq L(x,\bar{\lambda},\bar{\zeta},\bar{z}) \\ &\geq L(\bar{x},\bar{\lambda},\bar{\zeta},\bar{z}) \\ &= \sup_{(\zeta,z)} L(\bar{x},\bar{\lambda},\zeta,z) \\ &= < \bar{\lambda}, F(\bar{x}) > . \end{aligned}$$

Hence, $L(\bar{x}, \bar{\lambda}, \bar{\zeta}, \bar{z}) = p_{\bar{\lambda}}$ and $\bar{x}$ is a p.p. minimizer of $(\mathcal{P})$. Furthermore,

$$d_{\bar{\lambda}} = \inf_x L(x, \bar{\lambda}, \bar{\zeta}, \bar{z}) = L(\bar{x}, \bar{\lambda}, \bar{\zeta}, \bar{z}) = p_{\bar{\lambda}}.$$

By the weak duality $d_{\bar{\lambda}} = L(\bar{x}, \bar{\lambda}, \bar{\zeta}, \bar{z})$, and (9) shows that such y belong to W and are p.p. maxima of $(\mathcal{D})$.

(ii) For the given $\bar{x}$, $\bar{\lambda}$, we have by (7)

$$\begin{aligned}
< \bar{\lambda}, F(\bar{x}) > &= p_{\bar{\lambda}} \\
&\leq \inf_{x \in X} L(x, \bar{\lambda}, \bar{\zeta}, \bar{z}) \\
&\leq L(\bar{x}, \bar{\lambda}, \bar{\zeta}, \bar{z}) \\
&\leq \sup_{(\zeta, z)} L(\bar{x}, \bar{\lambda}, \zeta, z) \\
&= < \bar{\lambda}, F(\bar{x}) > .
\end{aligned}$$

This means that the inequalities are in fact equalities, and therefore (8) holds. □

For strict converse duality results we need convexity conditions, as follows.

Theorem 7 *Let K be pointed, $(\mathcal{P})$ be convex and the attainable set V of $(\mathcal{P})$ be closed. Let, for any $\lambda \in K^{\bullet}$, $(\mathcal{P}_\lambda)$ be inf-stable and satisfy the q.g.c. Then, every p.p. maximum $\bar{y}$ of $(\mathcal{D})$ is also a p.p. minimum of $(\mathcal{P})$.*

Proof - We claim first that $\bar{y} \in V + K$. Indeed, suppose the contrary. Then, since $V + K$ is closed and convex, $\bar{y}$ can be separated from $V + K$ by $\lambda_1 \in K^* \backslash \{0\}$, and by Theorem 3 we have

$$\begin{aligned}
< \lambda_1, \bar{y} > &< \inf_{y \in V+K} < \lambda_1, y > \\
&= \inf_{y \in V} < \lambda_1, y > \\
&= \inf_x \sup_{(\zeta, z)} L(x, \lambda_1, \zeta, z) \\
&= \sup_{(\zeta, z)} \inf_x L(x, \lambda_1, \zeta, z).
\end{aligned}$$

(Note that Theorem 3 is still true for any $\lambda \in K^*\backslash\{0\}$, not only for $\lambda \in K^\bullet$.)

On the other hand, since $\bar{y} \in W$, $\lambda_2 \in K^\bullet$ exists such that

$$< \lambda_2, \bar{y} > \leq \inf_{y \in V} < \lambda_2, y > . \tag{10}$$

Hence, for $\lambda_\alpha := (1-\alpha)\lambda_1 + \alpha\lambda_2 \in K^\bullet$ with any $\alpha \in (0,1)$,

$$< \lambda_\alpha, \bar{y} > < \inf_{y \in V} < \lambda_\alpha, y > = \sup_{(\zeta,z)} \inf_x L(x, \lambda_\alpha, \zeta, z).$$

Consequently, one can find $(\bar{\zeta}, \bar{z}) \in \mathbb{R}_+ \times Z$ such that

$$< \lambda_\alpha, \bar{y} > \; < \; \inf_x L(x, \lambda_\alpha, \bar{\zeta}, \bar{z}).$$

This yields to the existence of such a $y' \in W_1$ that $y' \in \bar{y} + K\backslash\{0\}$, contradicting the maximality of $\bar{y}$. Thus $\bar{y} \in V + K$. Now suppose $\bar{y} \notin V$ and $\bar{y} = y_0 + k$ with $y_0 \in V$ and $0 \neq k \in K$. Then, by the pointedness of K, $< \lambda_2, y_0 > \; < \; < \lambda_2, \bar{y} >$. This contradiction to (10) shows that $\bar{y} \in V \cap W$. By Corollary 4 we are done. □

Example (Convexity is essential) : Let $X = \mathbb{R}$, $Y = Z = \mathbb{R}^2$. Take $K = M = \mathbb{R}^2_+$, $G(x) \equiv 0$, $F_2(x) = x$ and

$$F_1(x) = \begin{cases} x & \text{if} \quad x \geq 0, \\ \sqrt{-x} & \text{if} \quad -1 < x < 0, \\ -x & \text{if} \quad x \leq -1. \end{cases}$$

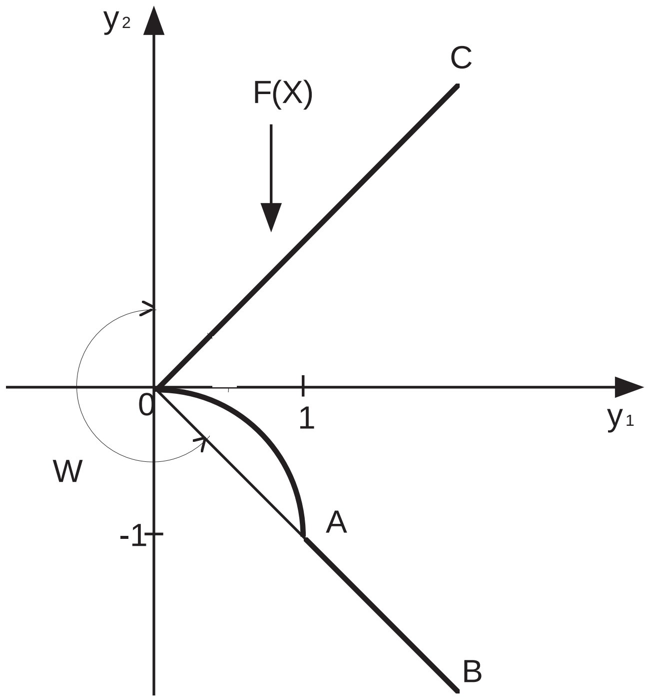
Then, Problem $(\mathcal{P})$ is unconstrained and easy to be solved. However using our augmented Lagrangian the dual $(\mathcal{D})$ can be still defined. A simple calculation shows that

$$\min_x \max_{(\zeta,z)} L(x,\lambda,\zeta,z) \;=\; \max_{(\zeta,z)} \min_x \; L(x,\lambda,\zeta,z)$$

$$= \begin{cases} 0 & \text{if} \quad \lambda_1 \geq \lambda_2, \\ -\infty & \text{if} \quad \lambda_1 < \lambda_2 \end{cases}$$

where $(\lambda_1, \lambda_2) \in K^\bullet = \text{Int } \mathbb{R}^2_+$.

Elementary argument shows that all points of half-straightline AB and $(0,0)$ are p.p. minima of $(\mathcal{P})$. p.p. maxima of $(\mathcal{D})$ consist of all points of the half-straightline OB. So the p.p.maxima of $(\mathcal{D})$ on open interval OA are not p.p. minima of $(\mathcal{P})$.



References

[1] Bertsekas D.P. (1982) *Constrained optimization and Lagrange multiplier methods*, Academic Press, New York.

[2] Boukari D. and Fiacco A.V. (1995) Survey of penalty, exact-penalty and multiplier methods from 1968 to 1993, Optimization, **32**, 301–334.

[3] Conn A.R., Gould N., Sartenaer A. and Toint P.L. (1996) Convergence properties of an augmented Lagrangian algorithm for optimization with a combination of general equality and linear constraints, SIAM J. Optim., **6**, 674–703.

[4] Craven B.D. (1989) A modified Wolfe dual for weak vector optimization, Numer. Funct. Anal. Optim., **10**, 899–907.

[5] Craven B.D. and Glover B.M. (1985) Invex functions and duality, J. Australian Math. Soc., **39** A, 1–20.

[6] Dolecki S. and Kurcyusz S. (1978) On ϕ-convexity in extremal problems, SIAM J. Control Optim., **16**, 277–300.

[7] Dolecki S. and Malivert C. (1993) General duality in vector optimization, Optimization, **27**, 97–119.

[8] Holmes R.B. (1972) *A course on optimization and best approximation*, Lec. Notes Math. 257, Springer, Berlin.

[9] Ito K. and Kunisch K. (1990) The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces, Math. Prog., **46**, 341–360.

[10] Ito K. and Kunisch K. (1996) Augmented Lagrangian SQP methods in Hilbert spaces and application to control in the coefficients problems, SIAM J. Optim., **6**, 96–125.

[11] Khanh P.Q. (1995) Invex-convexlike functions and duality, J.O.T.A., **87**, 141–165.

[12] Khanh P.Q. (1995) Sufficient optimality conditions and duality in vector opitmization with invex-convexlike functions, J.O.T.A., **87**, 359–378.

[13] Kiwiel K.C. (1996) On the twice differentiable cubic augmented Lagrangian, J.O.T.A., **88**, 233–236.

[14] Luc D.T. (1989) *Theory of vector optimization*, Lec. Notes Economics Math. Systems, Springer, Berlin.

[15] Pini R. and Singh C. (1997) A survey of recent [1985 – 1995] advances in generalized convexity with applications to duality theory and optimality conditions, Optimization, **39**, 311–360.

[16] Rockafellar R.T. (1974) Augmented Lagrange multiplier functions and duality in nonconvex programming, SIAM J. Control, **12**, 268–285.

[17] Sach P.H. and Craven B.D. (1991) Invex multifunctions and duality, Numer Funct. Anal. Optim., **12**, 575–591.

[18] Singh C., Bhatia D. and Rueda N. (1996) Duality in nonlinear multiobjective programming, J.O.T.A., **88**, 659–670.

[19] Song W. (1997) Lagrangian duality for minimization of nonconvex multifunctions, J.O.T.A., 93, 167–182.

[20] Weir T. and Mond B. (1989), Generalized convexity and duality in multiple-objective programming, Bull. Australian Math. Soc., **39**, 287–299.

[21] Wierzbicki A.P. and Kurcyusz S. (1977) Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in Hilbert space, SIAM J. Control Optim., **15**, 25–56.