JOHN F. TOLAND

On subdifferential calculus and duality in non-convex optimization

Mémoires de la S. M. F., tome 60 (1979), p. 177-183
<http://www.numdam.org/item?id=MSMF_1979__60__177_0>

Analyse non convexe [*1977. Pau*]
Bull. Soc. math. France,
Mémoire 60, 1979, p. 177-183.

ON SUBDIFFERENTIAL CALCULUS AND DUALITY IN NON-CONVEX OPTIMIZATION

John F. TOLAND

The purpose of this lecture is to give a brief résumé in an abstract setting of some results in the theory of duality for non-convex optimisation which have been obtained by the author recently. These results were initially motivated by a problem arising in mechanics (treated in [1] and [3]) but appear to have a larger field of application. This lecture does not contain the applications, for which we refer the reader to the references [1], [2], [3].

§.1 THE GENERAL DUALITY PRINCIPLE

Let V and V^* be linear topological vector spaces in separating duality, and let $< , > : V \times V^* \to \mathbb{R}$ denote the bilinear form which determines the duality between V and V^*.

We shall consider the extremality problem

$$\mathcal{P} \qquad \inf_{u \in V} \quad J(u)$$

As this stage we do not make any assumptions about the behaviour of J, and so we might just as well consider the maximisation problem

$$\sup_{u \in V} \quad - J(u)$$

instead.

Now let Y, Y^* be another pair of vector spaces in duality and let $\Phi : V\times Y \to \overline{\mathbb{R}}$ be a functional with the following properties : $\Phi(u,0) = - J(u)$ for all $u \in V$ and for each $u \in V$ the mapping $\Phi_u : Y \to \mathbb{R}$ defined by $\Phi_u(p) = \Phi(u,p)$ all $p \in Y$ is either convex and lower semi-continuous on Y, or has the property that $\Phi_u^{**}(0) = \Phi_u(0)$ for all $u \in V$.

This assumption is enough to ensure that an extremal principle, equivalent to $\mathcal{P}$ can be defined on Y^*, in much the same way as Legendre and Hamilton proceeded originally.

We begin by defining the Lagrangian functional $\mathcal{L} : V \times Y^* \to \overline{\mathbb{R}}$ as

$$- \mathcal{L}(u,p^*) = \sup_{p \in Y} \{ < p,p^* > - \Phi (u,p) \} .$$

Now put
$$-L\,p^* = \sup_{u\in V} \mathcal{L}(u,p^*)$$

Then the dual optimisation problem is

$$\mathcal{P}^* \qquad \inf_{p^*\in Y^*} L(p^*)$$

THEOREM 1 If $\Phi : V \times Y \to \overline{\mathbb{R}}$ is any functional such that $-\Phi(u,0) = Ju$, then

$$\inf_{u\in V} Ju \leqslant \inf_{p^*\in Y^*} L\,p^*$$

Proof :
$$-\inf_{u\in V} Ju = \sup_{u\in V} \Phi_u(0) \geqslant \sup_{u\in V} \Phi_u^{**}(0)$$

$$= \sup_{u\in V} \sup_{p^*\in Y^*} -\Phi_u^*(p^*) = \sup_{p^*\in Y^*} \sup_{u\in V} \mathcal{L}(u,p^*)$$

$$= \sup_{p^*\in Y^*} -Lp^* = -\inf_{p^*\in Y^*} Lp^* \qquad \text{Q.E.D.}$$

THEOREM 2 If $\Phi_u^{**}(0) = \Phi_u(0)$ for all $u \in V$ then

$$\inf_{u\in V} Ju = \inf_{p^*\in V^*} L\,p^* .$$

Proof : Immediate from previous result.

Remarks : We have used the Lagrangian $\mathcal{L}$ to define a dual extremal problem, and not to define a minimax problem equivalent to $\mathcal{P}$. The existence of a saddle-point $(\underline{u},\underline{p}^*)$ for $\mathcal{L}$ ($\mathcal{L}(\underline{u},p^*) \leqslant \mathcal{L}(\underline{u},\underline{p}^*) \leqslant \mathcal{L}(u,\underline{p}^*)$) implies that $\underline{u}$ is a solution of the problem $J\underline{u} = \sup_{u\varepsilon V} Ju$. So if J is not bounded above $\mathcal{L}$ does not have a saddle-point, but nonetheless it can be used to define the dual extremal problem $\mathcal{P}^*$ as we have shown.

The duality result in theorem 2 is true whether solutions for $\mathcal{P}$ or $\mathcal{P}^*$ exist or not. In our next result we examine the relationship between solutions of $\mathcal{P}$ and of $\mathcal{P}^*$.

THEOREM 3 Let $\Phi : V \times V^* \to \overline{\mathbb{R}}$ be such that $-\Phi(u,0) = Ju$ for all $u \varepsilon V$. Suppose that $\underline{u}$ is a solution of $\mathcal{P}$ and that $\underline{p}^* \in \partial\Phi_{\underline{u}}(0)$. Then $\underline{p}^*$ solves $\mathcal{P}^*$ and

$$\left.\begin{aligned} -\mathcal{L}(\underline{u},\underline{p}^*) + \Phi(\underline{u},0) &= 0 \quad &\text{(a)} \\ L(p^*) + \mathcal{L}(\underline{u},\underline{p}^*) &= 0 \quad &\text{(b)} \end{aligned}\right\} \qquad (E)$$

Proof : Because of theorem 1 it will suffice to show that $L(\underline{p}^*) = J(\underline{u})$.

$$\begin{aligned} -\mathcal{L}(\underline{u},\underline{p}^*) &= \sup_{p\in V} \{ \langle p,\underline{p}^* \rangle - \Phi(\underline{u},p)\} \\ &\leq \sup_{p\in V} \{ \langle p,\underline{p}^* \rangle - \{\Phi(\underline{u},0) + \langle p,\underline{p}^* \rangle \}\} \\ &= -\Phi(\underline{u},0) . \end{aligned}$$

Hence $\quad -L(\underline{p}^*) = \sup_{u\in V} \mathcal{L}(\underline{u},\underline{p}^*) \geq \mathcal{L}(\underline{u},\underline{p}^*) \geq \Phi(\underline{u},0) = J(\underline{u})$

and so $\quad L(\underline{p}^*) = J\underline{u}$.

By theorem 1 , $L(\underline{p}^*)= J(\underline{u})$, and it is clear that

$$-\mathcal{L}(\underline{u},\underline{p}^*) + \Phi(\underline{u},0) = 0$$

$$L(\underline{p}^*) + \mathcal{L}(\underline{u},\underline{p}^*) = 0 \qquad\qquad \text{Q.E.D.}$$

Remark : It is apparent that these extremality conditions comprise the Euler-Lagrange equations when the extremal problem arises in the calculus of variations (see [1] , [2] for further details). When we treat the important special case of J in the form G - F with F convex in the next section we will see that they are not enough to ensure that $\underline{u}$ and $\underline{p}^*$ solve $\mathcal{P}$ and $\mathcal{P}^*$ respectively. However the following result is true. If $\underline{p}^* \in Y^*$ and $\{u_n\} \subset V$ is such that

$$\lim_{n\to\infty} \mathcal{L}(u_n,\underline{p}^*) = -L(\underline{p}^*) .$$

Then $\underline{p}^*$ solves $\mathcal{P}^*$ if and only if $\{u_n\}$ is a minimising sequence for $\mathcal{P}$ and $\Phi(u_n,0) - \mathcal{L}(u_n,\underline{p}^*) \to 0$ as $n \to \infty$.

§.2. AN IMPORTANT SPECIAL CASE

In order to simplify our discussion further we shall give an analysis of the special case which lead us to the results of the previous section and of [2] .

Suppose $Ju = Gu - Fu$ for all $u \in V$. Then if we put

$$\Phi(u,p) = F(u + p) - Gu , \text{ for all } (u,p) \in V \times V$$

$$\mathcal{L}(u,p^*) = \langle u,p^* \rangle - Gu - F^* p^*$$

and

$$L(p^*) = F^*p^* - G^* p^* \ .$$

In this context theorems 1, 2 and 3 take the following more specific form :

THEOREM 1' : $\inf_{u\in V} G(u) - F(u) \leq \inf_{p^* \in V^*} F^*(p^*) - G^*(p^*)$.

THEOREM 2' : If F is convex and lower semi-continuous on V , then

$$\inf_{u\in V} Gu - Fu = \inf_{p^* \in V^*} F^*p^* - G^*p^*$$

THEOREM 3' : If $\underline{u} \in V$ is such that

$$G(\underline{u}) - F(\underline{u}) = \inf_{u\in V} G(u) - F(u)$$

and $\underline{p}^* \in \partial F(\underline{u})$, then

$$F^*\underline{p}^* - G^*\underline{p}^* = \inf_{p^* \in Y^*} F^*p^* - G^*p^*$$

$$\left.\begin{aligned} F\underline{u} + F^*\underline{p}^* &= \langle \underline{u},\underline{p}^* \rangle \quad &(a) \\ G\underline{u} + G^*\underline{p}^* &= \langle \underline{u},\underline{p}^* \rangle \quad &(b) \end{aligned}\right\} \quad (E')$$

Remark : It is clear in this special case why the extremality conditions E, E' of theorems 3, 3' are not enough to ensure that a pair $(\underline{u},\underline{p}^*)$ are solutions of $\mathcal{P}$ and $\mathcal{P}^*$ respectively. In fact E' holds if and only if

$$\underline{p}^* \in \partial G(\underline{u}) \cap \partial F(\underline{u}) \ .$$

We can make the following definition. A point $u \in V$ is a critical point of $G-F$ if $\partial G(u) \cap \partial F(u) \ni p^*$ for some $p^* \in V^*$.

It is then a natural question to ask if E' is satisfied by a pair $(\underline{u},\underline{p}^*) \in V \times V$, under what circumstances can we be sure that $\underline{u}$ is a minimiser for $\mathcal{P}$. In keeping with our approach throughout this lecture we want the answer in terms of subdifferentials.

THEOREM 4 Suppose G and F are convex functionals which are such that ∂F and ∂G are single-valued and hemi-continuous mappings from V into V^* . Suppose that $\underline{u} \in V$ is a critical point of $G - F$ in the sense which we have just defined (i.e. $\partial G(\underline{u}) = \partial F(\underline{u})$).

If

$$<\partial G(u),\, u-\underline{u}> \;\geqslant\; <\partial F(u), u-\underline{u}>$$

for all u in a neighbourhood N of $\underline{u}$ in V, then $G(\underline{u})-F(\underline{u}) \leqslant G(u)-F(u)$ for all $u \in N$, and $\underline{u}$ is a local minimiser of $G - F$.

Proof : Let $u \in N$; define the mapping

$$h(t) = G(tu + (1-t)\underline{u}) - F(tu + (1-t)\underline{u}) \quad .$$

Now $h : [0,1] \to \mathbb{R}$ is continuous and we want to show that $h(1) \geqslant h(0)$. By a classical result due to Borel it will suffice to show that the lower symmetric derivative of h is non-negative at each point of $(0,1)$. In other words we must show that

$$\liminf_{s \to 0} \frac{h(t+s) - h(t-s)}{2s} \geqslant 0$$

for all $t \in (0,1)$.

But
$$G((t+s)u + (1-t-s)\underline{u}) - G((t-s)u + (1-t+s)\underline{u})$$

$$\geqslant \; < \partial G((t-s)u + (1-t+s)\underline{u}) \;,\; 2s(u-\underline{u}) >$$

and
$$F((t-s)u + (1-t+s)\underline{u}) - F((t+s)u + (1-t-s)\underline{u})$$

$$\geqslant \; < \partial F((t+s)u+(1-t-s)\underline{u}),\; -2s(u-\underline{u}) > \quad .$$

So

$$\frac{h(t+s)-h(t-s)}{2s} \;\geqslant\; < \partial G((t-s)u + (1-t+s)\underline{u}) - \partial F((t+s)u+(1-t-s)\underline{u},\; u-\underline{u} >$$

$$\longrightarrow \; < \partial G(tu + (1-t)\underline{u}) - \partial F(tu+(1-t)\underline{u}),\; u-\underline{u} > \;\geqslant 0 \quad .$$

Hence h is an increasing function on $(0,1)$ and the result is established. Q.E.D

Now it is obvious that there is a symmetry between problem $\wp$:

$$\inf_{u \in V} \; Gu - Fu$$

and problem $\wp^*$:

$$\inf_{p^* \in V^*} \; F^*p^* - G^*p^*$$

If $(\underline{u},\underline{p}^*)$ satisfy E' then we say that $\underline{u}$ is a critical point of $G - F$. By the same token p^* is a critical point of $F^* - G^*$, and we can apply theorem 4 to decide whether or not $\underline{p}^*$ is a local minimiser for $\wp^*$. But it is also an interesting question to ask whether we can deduce the local properties of $\underline{p}^*$ from a knowledge of the local properties of $\underline{u}$.

<u>THEOREM 5</u> Let $(\underline{u},\underline{p}^*) \in V \times V^*$ satisfy E' , and suppose that $\underline{u}$ is a local minimiser of $G - F$.

If ∂G^* is single-valued and continuous on V^* , then $\underline{p}^*$ is a local minimiser for $F^* - G^*$.

<u>Proof</u> : Since $\underline{u}$ is a local minimiser for $G - F$ there exists a neighbourhood N of u such that if $u \in N$ then

$$Gu - Fu \geqslant G\underline{u} - F\underline{u} \quad .$$

Now by the continuity of ∂G^* there exists a neighbourhood M of $\underline{p}^*$ such that if $p^* \in M$ then $\partial G^*(p^*) = u \in N$. Thus if $p^* \in M$, then for some $u \in N$

$$Gu + G^*p^* = \langle u,p^* \rangle$$

and

$$Fu + F^*p^* \geqslant \langle u,p^* \rangle$$

Thus

$$Gu - Fu \leqslant F^*p^* - G^*p^*$$

But

$$F^*\underline{p}^* - G^*\underline{p}^* = G\underline{u} - F\underline{u} \qquad \text{by E'} ,$$

and so

$$F^*p^* - G^*p^* \geqslant F^*\underline{p}^* - G^*\underline{p}^*$$

for all $p^* \in M$ and the proof is complete. Q.E.D.

So, in this lecture we have illustrated in an abstract setting how the subdifferential calculus may play a useful role in the problems of non-convex optimisation and in the calculus of variations. In a recent paper [1] the author has established the results on problems of the form $G - F$ quite independently of the considerations of section 1 of this lecture. There results on the duality of minimising sequences are established,and the theory is applied in the analysis of a problem in mechanics. In [2] the results of section 1 are analysed in greater detail than here, and their application to the calculus of variations is given. In particular the Euler-Lagrange equations are seen to hold under weak hypotheses on the integrands. In [3] the duality between local properties of critical points is developed in the spirit of theorem 5 above ; but a more sophisticated context is needed for the intended application to the stability question for the heavy chain which is also treated in that paper.

REFERENCES

[1] TOLAND J.F. : A duality principle for non-convex optimisation and the calculus of variations. F.M.R.I. (University of Essex) report N° 77 to appear Arch. Rational Mech. Analysis.

[2] TOLAND J.F. : Duality in non-convex optimisation. F.M.R.I. (University of Essex) report N° 78 to appear Jour. Math. Anal. Appl.

[3] TOLAND J.F. : On the stability of rotating heavy chains. F.M.R.I. (University of Essex) report N° 82 (May 1977).