Article

An Extremum Principle for Smooth Problems

Dariusz Idczak [1,*] **and Stanisław Walczak** [1,2]

[1]   Faculty of Mathematics and Computer Science, University of Lodz, Banacha 22, 90-238 Lodz, Poland; stanislaw.walczak@wmii.uni.lodz.pl

[2]   Faculty of Mathematics and Computer Science, Stefan Batory State University, Batorego 64C, 96-100 Skierniewice, Poland

*   Correspondence: dariusz.idczak@wmii.uni.lodz.pl

check for updates

Abstract: We derive an extremum principle. It can be treated as an intermediate result between the celebrated smooth-convex extremum principle due to Ioffe and Tikhomirov and the Dubovitskii–Milyutin theorem. The proof of this principle is based on a simple generalization of the Fermat's theorem, the smooth-convex extremum principle and the local implicit function theorem. An integro-differential example illustrating the new principle is presented.

Keywords: extremum principle; Fermat's theorem; local implicit function theorem

MSC: 90C48; 49K27

1. Introduction

Let us consider the following problem:
(P) minimize locally in both variables the function

$$f_0 : X \times Y \to \mathbb{R} \tag{1}$$

under constraints

$$f(x, u) = 0, \tag{2}$$

$$u \in U, \tag{3}$$

where $f : X \times Y \to Z$, X, Y, Z are real Banach spaces, $U \subset Y$ is a fixed set. By a local in (x, u) solution to this problem we mean a pair $(x_*, u_*) \in X \times Y$ satisfying the constraints (2), (3) and such that

$$f_0(x_*, u_*) \leq f_0(x, u) \tag{4}$$

for any pair $(x, u) \in V_{x_*} \times V_{u_*}$ satisfying the constraints, where V_{x_*} is a neighborhood of x_* in X and U_{u_*} is a neighborhood of u_* in Y.

The aim of the paper is to derive an extremum principle for the problem **(P)**, giving necessary conditions for its solution. Such conditions allow one to find pairs (x, u) suspected of being the solutions of the problem under consideration. More precisely, points that do not satisfy the necessary conditions cannot be the solutions.

There are two known main powerful tools giving the necessary conditions for problems of such a type. The first of them—the smooth-convex extremum principle due to Ioffe and Tikhomirov (see [1] (Part 1.1.2, Theorem 3) and also Theorem 2 below)—can be used to study local in x solutions to **(P)**. Besides the standard smoothness and regularity assumptions in x imposed on f_0 and f, it contains a "convexity" assumption imposed on these functions with respect to u. In this theorem, one does

not require the closedness of U nor non-emptiness of the interior of U. The second tool is the Dubovitskii–Milyutin theorem (see [2–4] for a systematic exposition). From this theorem, one can deduce necessary conditions for local in (x, u) solution to (**P**). In the formulation of this theorem, some cones and the corresponding conjugate cones appear. To use the useful characterizations of the cones, it must be assumed that f_0 and f are smooth with respect to (x, u) and the set U is closed and has a non-empty interior.

Let us point out that the paper [5] gives the most recent discussion of the smooth-convex extremum principle. In particular, an extension of this principle—Lagrange's principle for smoothly approximately convex problems—has been derived for a problem containing some additional "membership" constraint of type $G(x) \in Q$. The main novelty of this theorem lies in replacing the smoothness and convexity assumptions by a smoothly approximate convexity assumption. The paper also includes a very interesting historical commentary with relevant references.

Our principle (Theorem 3) gives necessary conditions for local in (x, u) solution to (**P**) under smoothness of f_0 and f in (x, u), without convexity and approximate convexity assumptions imposed on f and f_0 (we assume only the convexity of the set U) and without any assumptions on the closedness of U as well as on the interior of U. Proof of this result is short and based on a very simple generalization of the Fermat's theorem, the smooth-convex principle applied to the linear problem (**P**) and the local implicit function theorem. An example illustrating the obtained result is presented. It shows that using the new principle one can improve the maximum principle derived in [6] with the aid of the Dubovitskii–Milyutin theorem.

2. Preliminaries

In this section, we recall a generalization of the Fermat's theorem, implicit function theorem and a particular case of the smooth-convex extremum principle.

We say that a function $g : U \to \mathbb{R}$ where U is a subset of a real Banach space Y, has a directional derivative at $u \in U$ in a direction $h \in Y$, if there exists $t_0 > 0$ such that $u + th \in U$ for $t \in (0, t_0)$ and the limit

$$\lim_{t \to 0^+} \frac{g(u + th) - g(u)}{t}$$

exists. In such a case, this limit is denoted as $g'_h(u)$ and called the directional derivative of g at u in the direction h. We have the following lemma generalizing Fermat's theorem (proof of this lemma is immediate).

Lemma 1. *Let U be a subset of a real Banach space Y. If a function $g : U \to \mathbb{R}$ has the directional derivative at $u_* \in U$ in the direction $h \in Y$ and u_* is a local minimum point of g on U, then*

$$0 \le g'_h(u_*). \tag{5}$$

The classical local implicit function theorem in Banach spaces is the following theorem (see [1]).

Theorem 1 (local implicit function theorem). *Let X, Y, Z be real Banach spaces, V—a neighborhood of a point (x_0, y_0) in $X \times Y$ and $F : V \to Z$—a mapping of class C^1. Assume that $F(x_0, y_0) = 0$ and the partial differential $F_x(x_0, y_0) : X \to Z$ is bijective. Then, there exist balls $B(x_0, \varepsilon)$, $B(y_0, \delta)$ and a mapping $\lambda : B(y_0, \delta) \to B(x_0, \varepsilon)$ such that*

- *equalities $F(x, y) = 0$ and $x = \lambda(y)$ are equivalent in the set $B(x_0, \varepsilon) \times B(y_0, \delta)$*
- *λ is of class C^1 and*

$$\lambda'(y) = -[F_x(\lambda(y), y))]^{-1} \circ F_y(\lambda(y), y)$$

for any $y \in B(y_0, \delta)$.

Now, let us consider the following problem:

(**S**) minimize locally in x the function

$$g_0 : X \times U \to \mathbb{R}$$

under constraints

$$g(x, u) = 0, \tag{6}$$

where $g : X \times U \to Z$, X, Z are real Banach spaces, U is any fixed set. By a local in x solution to this problem we mean a pair $(x_*, u_*) \in X \times U$ satisfying the constraints (6) and such that

$$g_0(x_*, u_*) \leq g_0(x, u) \tag{7}$$

for any pair $(x, u) \in V_{x_*} \times U$ satisfying (6), where V_{x_*} is a neighborhood of x_*.

A particular case of the smooth-convex extremum principle (see [1]) is the following theorem.

Theorem 2. *Let (x_*, u_*) be a local in x solution to the problem (**S**).*

- *for any $u \in U$, the mappings*

$$X \ni x \mapsto g_0(x, u) \in \mathbb{R},$$

$$X \ni x \mapsto g(x, u) \in Y,$$

  are of class C^1 at x_*,
- *for any $x \in V_{x_*}$ where V_{x_*} is a neighborhood of x_* the mappings*

$$U \ni u \mapsto g_0(x, u) \in \mathbb{R},$$

$$U \ni u \mapsto g(x, u) \in Y,$$

  satisfy the following „convexity" assumption: for any $u_1, u_2 \in U$, $\alpha \in (0, 1)$ there exists $u \in U$ such that

$$g_0(x, u) \leq \alpha g_0(x, u_1) + (1 - \alpha)g_0(x, u_2),$$

$$g(x, u) = \alpha g(x, u_1) + (1 - \alpha)g(x, u_2),$$

- *the differential*

$$g_x(x_*, u_*) : X \to Y$$

  is onto,

then there exists $\mu \in Y^*$ (conjugate space ([1])) such that

$$(g_0)_x(x_*, u_*)x + \langle \mu, (g_x(x_*, u_*)x) \rangle = 0$$

for any $x \in X$ and

$$g_0(x_*, u_*) + \langle \mu, g(x_*, u_*) \rangle = \min_{u \in U}(g_0(x_*, u) + \langle \mu, g(x_*, u) \rangle).$$

---

[1]   By conjugate space (dual space) Y^* to a real Banach space Y we mean the space of all linear continuous functionals $\mu : Y \to \mathbb{R}$.

3. An Extremum Principle

Assume that a point $(x_*, u_*) \in X \times Y$ is a local in (x, u) minimum point for the problem (**P**) with a set $U \subset Y$. Moreover, assume that

1.　$f_0 : X \times Y \to \mathbb{R}$ is Frechet differentiable at (x_*, u_*)
2.　$f : X \times Y \to Z$ is of class C^1 on some neighborhood of (x_*, u_*)
3.　$f_x(x_*, u_*) : X \to Z$ is bijective.

From the local implicit function theorem applied to f, it follows that there exist balls $B(u_*, \delta)$ and $B(x_*, \varepsilon)$ and a mapping $\lambda : B(u_*, \delta) \to B(x_*, \varepsilon)$ of class C^1 with differential

$$\lambda'(u) = -[f_x(\lambda(u), u)]^{-1} \circ f_u(\lambda(u), u),$$

such that

$$f(\lambda(u), u) = 0$$

for $u \in B(u_*, \delta)$ ($\lambda(u)$ is the unique point in $B(x_*, \varepsilon)$ such that the last equality holds true).

Consider the mapping

$$\widetilde{g} : B(u_*, \delta) \ni u \longmapsto (\lambda(u), u) \longmapsto f_0(\lambda(u), u) \in \mathbb{R}$$

Of course, this mapping is differentiable in u_* and the differential $\widetilde{g}'(u_*) : Y \to \mathbb{R}$ at u_* is of the form

$$\widetilde{g}'(u_*)u = (f_0)_x(x_*, u_*)\lambda'(u_*)u + (f_0)_u(x_*, u_*)u$$
$$= -(f_0)_x(x_*, u_*)([f_x(x_*, u_*)]^{-1}(f_u(x_*, u_*)u)) + (f_0)_u(x_*, u_*)u$$

for $u \in Y$.

Now, let us assume that the set $U \subset Y$ is convex and consider the mapping $g = \widetilde{g} \mid_{B(u_*, \delta) \cap U}$. Since U is convex and $\widetilde{g}$ is differentiable at u_*, g has the directional derivative at u_* in any direction $h = u - u_*$ with $u \in U$. Clearly,

$$g'_{u-u_*}(u_*) = \widetilde{g}'(u_*)(u - u_*) = \widetilde{g}'(u_*)u - \widetilde{g}'(u_*)u_*.$$

It is easy to observe that u_* is the local minimum point of g. So, from Lemma 1, it follows that

$$\widetilde{g}'(u_*)u_* \le \widetilde{g}'(u_*)u$$

for any $u \in U$, i.e.,

$$- (f_0)_x(x_*, u_*)([f_x(x_*, u_*)]^{-1}(f_u(x_*, u_*)u_*)) + (f_0)_u(x_*, u_*)u_*$$
$$\le -(f_0)_x(x_*, u_*)([f_x(x_*, u_*)]^{-1}(f_u(x_*, u_*)u)) + (f_0)_u(x_*, u_*)u$$

for $u \in U$. Denoting

$$w = [f_x(x_*, u_*)]^{-1}(f_u(x_*, u_*)u),$$

$$w_* = [f_x(x_*, u_*)]^{-1}(f_u(x_*, u_*)u_*)$$

we see that

$$(f_0)_x(x_*, u_*)(-w_*) + (f_0)_u(x_*, u_*)u_* \le (f_0)_x(x_*, u_*)(-w) + (f_0)_u(x_*, u_*)u$$

for any $u \in U$. Clearly,

$$f_x(x_*, u_*)(-w) + f_u(x_*, u_*)u = 0,$$

$$f_x(x_*, u_*)(-w_*) + f_u(x_*, u_*)u_* = 0.$$

In other words, the pair $(z_*, u_*) \in X \times U$ with $z_* = -w_*$ is a solution to problem:

(LIN) minimize globally in (z, u) the function

$$h_0 : X \times U \to \mathbb{R}$$

given by

$$h_0(z, u) = (f_0)_x(x_*, u_*)z + (f_0)_u(x_*, u_*)u$$

under constraints

$$h(z, u) = 0$$

where

$$h(z, u) = f_x(x_*, u_*)z + f_u(x_*, u_*)u.$$

Linearity of the mappings h_0, h and regularity of the differential $f_x(x_*, u_*)$ imply that all assumptions of Theorem 2 are satisfied for the problem (**LIN**). Consequently, there exists $\mu \in Y^*$ such that

$$(f_0)_x(x_*, u_*)x + \langle \mu, f_x(x_*, u_*)x \rangle = 0 \tag{8}$$

for any $x \in X$ and

$$(f_0)_u(x_*, u_*)u_* + \langle \mu, f_u(x_*, u_*)u_* \rangle = \min_{u \in U}((f_0)_u(x_*, u_*)u + \langle \mu, f_u(x_*, u_*)u \rangle). \tag{9}$$

Thus, we have proven the following extremum principle.

Theorem 3. *If $(x_*, u_*) \in X \times Y$ is a local in (x, u) minimum point for the problem (P) with a convex set $U \subset Y$ and assumptions (1.)–(3.) are satisfied, then there exists $\mu \in Y^*$ such that (8) and (9) hold true.*

4. An Application

In paper [6], we consider the following optimal control problem described by the nonlinear integro-differential equation of Volterra type

$$\begin{cases} x'(t) + \int_a^t \Phi(t, \tau, x(\tau), u(\tau))d\tau = \Psi(t, x(t), v(t)), \ t \in J := [a, b] \text{ a.e.,} \\ x(a) = 0 \end{cases} \tag{10}$$

with constraints

$$u \in M, \ v \in N \tag{11}$$

and the nonlinear performance index of Bolza type

$$f_0(x, u) = \int_a^b F_0(t, x(t), u(t), v(t))dt + G_0(x(b)) \tag{12}$$

where $\Phi : P_\Delta \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $\Psi : J \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$, $F_0 : J \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}$, $G_0 : \mathbb{R}^n \to \mathbb{R}$, $M \subset \mathbb{R}^m, N \subset \mathbb{R}^r$ and

$$P_\Delta = \{(t, \tau) \in J \times J; \ \tau \leq t\}.$$

Consider the problem:
(ID) minimize locally in both variables (x and (u, v)) the function

$$f_0(x, (u, v))$$

under constraints

$$f(x,(u,v)) = 0,$$

$$(u,v) \in U = U_1 \times U_2,$$

where

$$f_0 : AC_0^2 \times L_{m+r}^\infty \ni (x,(u,v)) \longmapsto \int_a^b F_0(t,x(t),u(t),v(t))dt + G_0(x(b)) \in \mathbb{R},$$

$$f : AC_0^2 \times L_{m+r}^\infty \ni (x,(u,v)) \longmapsto x'(t) + \int_a^t \Phi(t,\tau,x(\tau),u(\tau))d\tau - \Psi(t,x(t),v(t)) \in L^2,$$

with the set of solutions $AC_0^2 = AC_0^2(J,\mathbb{R}^n)$ (set of absolutely continuous functions possessing squared integrable derivatives, vanishing at $t = 0$) and the set of functional parameters (controls) $U = U_1 \times U_2$ with

$$U_1 = L_m^\infty(J,M) \subset L_m^\infty = L_m^\infty(J,\mathbb{R}^m),$$

$$U_2 = L_r^\infty(J,N) \subset L_r^\infty = L_r^\infty(J,\mathbb{R}^r)$$

consisting of essentially bounded functions taking their values in the sets $M \subset \mathbb{R}^m$, $N \subset \mathbb{R}^r$, respectively. On the sets M, N, we assume that they are convex.

Checking differentiability of f_0, f and regularity of f just as in [6], we can obtain the maximum principle given in [6] (Theorem 4.1) not assuming that the sets M, N are closed with nonempty interiors (it is sufficient to assume only convexity of these sets).

5. Conclusions

In the paper, we derive a new extremum principle. In this principle, we do not impose any convexity nor approximate convexity assumption on the functions f_0, f describing the problem (**P**)—such assumptions appear in the smooth-convex extremum principle and in the Lagrange's principle for smoothly approximately convex problems. We also do not assume that the set U is closed and has nonempty interior—such an assumption usually appears when we apply the Dubovitskii–Milyutin theorem to the problem (**P**) (see [6]).

References

1. Ioffe, A.D.; Tikhomirov, V.M. *Theory of Extremal Problems*; Elsevier: Amsterdam, The Netherlands, 1979.
2. Dubovitskii, M.Y.; Milyutin, A.A. The extremum problem in the presence of constraints. *Dokl. Acad. Nauk SSSR* **1963**, *149*, 759–762.
3. Dubovitskii, M.Y.; Milyutin, A.A. Extremum problems in the presence of constraints. *Zh. Vychisl. Mat. Mat. Fiz.* **1965**, *5*, 395–453.
4. Girsanov, I.W. *Lectures on Mathematical Theory of Extremum Problems*; Springer: New York, NY, USA, 1972.
5. Avakov, E.R.; Magaril-Il'yaev, G.G.; Tikhomirov, V.M. Lagrange's principle in extremum problems with constraints. *Russ. Math. Surv.* **2013**, *68*, 401–433. [CrossRef]
6. Idczak, D.; Walczak, S. Necessary optimality conditions for an integro-differential Bolza problem via Dubovitskii-Milyutin method. *Discret. Contin. Dyn. Syst. B* **2019**, *24*, 2281. [CrossRef]