Journal of Nonlinear and Convex Analysis
Volume 10, Number 1, 2009, 149–156

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

EXACT PENALTY IN CONSTRAINED OPTIMIZATION AND CRITICAL POINTS OF LIPSCHITZ FUNCTIONS

ALEXANDER J. ZASLAVSKI

Abstract. In this paper we use the penalty approach to study two constrained minimization problems in infinite-dimensional Asplund spaces. A penalty function is said to have the exact penalty property if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. We establish a simple sufficient condition for exact penalty property using the notion of the Mordukhovich basic subdifferential.

1. Introduction

Penalty methods are an important and useful tool in constrained optimization. See, for example, [4-8, 11, 14, 18, 19, 21] and the references mentioned there. In this paper we use the penalty approach in order to study two constrained nonconvex minimization problems with Lipschitzian (on bounded sets) objective functions. The first problem is an equality-constrained problem in an Asplund space with a locally Lipschitzian constraint function and the second problem is an inequality-constrained problem in an Asplund space with a locally Lipschitzian constraint function. Note that a Banach space is an Asplund space if and only if every separable subspace has a separable dual [13].

A penalty function is said to have the exact penalty property if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. The notion of exact penalization was introduced by Eremin [10] and Zangwill [19] for use in the development of algorithms for nonlinear constrained optimization. For a detailed historical review of the literature on exact penalization see [4, 6, 8].

In [21] it was established the existence of a penalty coefficient for which approximate solutions of the unconstrained penalized problem are close enough to approximate solutions of the corresponding constrained problem. This is a novel approach in the penalty type methods.

Consider a minimization problem $h(z) \to \min$, $z \in X$ where $h : X \to R^1$ is a lower semicontinuous bounded from below function on a Banach space X. If the space X is infinite-dimensional or if the function h does not satisfy a coercivity assumption, then the existence of solutions of the problem is not guaranteed and in this situation we consider δ-approximate solutions. Namely, $x \in X$ is a δ-approximate solution of the problem $h(z) \to \min$, $z \in X$, where $\delta > 0$, if $h(x) \le \inf\{h(z) : z \in X\} + \delta$.

2000 *Mathematics Subject Classification.* 49M30, 90C26, 90C30.
Key words and phrases. Approximate solution, critical point, Ekeland's variational principle, minimization problem, Mordukhovich basic subdifferential, penalty function .

In [21] and in this paper we consider minimization problems in a general Banach space and in a general Asplund space respectively. Therefore we are interested in approximate solutions of the unconstrained penalized problem and in approximate solutions of the corresponding constrained problem. Under certain mild assumptions we show the existence of a constant $\Lambda_0 > 0$ such that the following property holds:

For each $\epsilon > 0$ there exists $\delta(\epsilon) > 0$ which depends only on ϵ such that if x is a $\delta(\epsilon)$-approximate solution of the unconstrained penalized problem whose penalty coefficient is larger than Λ_0, then there exists an ϵ-approximate solution y of the corresponding constrained problem such that $||y - x|| \le \epsilon$.

This property implies that any exact solution of the unconstrained penalized problem whose penalty coefficient is larger than Λ_0, is an exact solution of the corresponding constrained problem. Indeed, let x be a solution of the unconstrained penalized problem whose penalty coefficient is larger than Λ_0. Then for any $\epsilon > 0$ the point x is also a $\delta(\epsilon)$-approximate solution of the same unconstrained penalized problem and in view of the property above there is an ϵ-approximate solution y_ϵ of the corresponding constrained problem such that $||x - y_\epsilon|| \le \epsilon$. Since ϵ is an arbitrary positive number we can easily deduce that x is an exact solution of the corresponding constrained problem. Therefore our results also include the classical penalty result as a special case.

In [21] the existence of the constant Λ_0 for the equality-constrained problem was established under the assumption that the set of admissible points does not contains critical points of the constraint function. The notion of critical points used in [21] is based on Clarke's generalized gradients [7]. It should be mentioned that there exists also the construction of Mordukhovich basic subdifferential introduced in [12] which is intensively used in the literature. See, for example, [13, 14] and the references mentioned there. In the present paper we generalize the results of [21] for minimization problems on Asplund spaces using the (less restrictive) notion of critical points via Mordukhovich basic subdifferential.

2. Main results

Let $(X, ||\cdot||)$ be an Asplund space and $(X^*, ||\cdot||_*)$ its dual equipped with the weak* topology w^*.

If $F : X \to 2^{X^*}$ is a set-valued mapping between the Banach space X and its dual X^*, then the notation

$$\limsup_{x \to \bar{x}} F(x) := \{x^* \in X^* : \text{ there exist sequences } x_k \to \bar{x} \text{ and } x_k^* \overset{w^*}{\to} \to x^*$$

$$\text{as } k \to \infty \text{ with } x_k^* \in F(x_k) \text{ for all natural numbers } k\} \tag{2.1}$$

signifies the sequential Painleve-Kuratowski upper limit with respect to the norm topology of X and the weak* topology of X^*.

For each $x^* \in X^*$ and each $r > 0$ set

$$B_*(x^*, r) = \{l \in X^* : ||l - x^*||_* \le r\}.$$

In this paper in order to obtain a sufficient condition for the existence of an exact penalty we use the notion of Mordukhovich basic subdifferential introduced in [12] (see also [13, page 82]). In order to meet this goal we first present the notion of an analytic ϵ-subdifferential (see [13, page 87]).

Let $\phi : X \to R^1$, $\epsilon > 0$ and let $\bar{x} \in X$. Then the set

(2.2) $\widehat{\partial}_{a\epsilon}\phi(\bar{x}) := \{x^* \in X^* : \liminf_{x\to\bar{x}}[(\phi(x) - \phi(\bar{x}) - < x^*, x - \bar{x} >)||x - \bar{x}||^{-1}] \geq -\epsilon\}$

is the analytic ϵ-subdifferential of ϕ at $\bar{x}$.

By Theorem 1.8.9 of [13, page 92], the set

$$\partial\phi(\bar{x}) = \limsup_{x\xrightarrow{\phi}\to\bar{x},\ \epsilon\to 0^+} \widehat{\partial}_{a\epsilon}\phi(x) \tag{2.3}$$

is Mordukhovich basic (limiting) subdifferential of the function ϕ at the point $\bar{x}$.

It should be mentioned that in view of Theorem 2.34 of [13, p. 218],

$$\partial\phi(\bar{x}) = \limsup_{x\xrightarrow{\phi}\to\bar{x}} \widehat{\partial}_{a0}\phi(x).$$

Here we use the notation that $x \xrightarrow{\phi} \to\bar{x}$ if and only if $x \to \bar{x}$ with $\phi(x) \to \phi(\bar{x})$, where $\phi(x) \to \phi(\bar{x})$ is superfluous if ϕ is continuous at $\bar{x}$.

Note that in the present paper we do not provide a definition of Mordukhovich basic (limiting) subdifferential as it appears in the literature [13, page 82]. Instead of it we work with the formula (2.3) which is more convenient for our goals.

Let $f : X \to R^1$ be a locally Lipschitzian function. For each $x \in X$ denote by $\partial f(x)$ Mordukhovich basic subdifferential of f at x and set

$$\Xi_f(x) = \inf\{||l||_* : \ l \in \partial f(x)\}. \tag{2.4}$$

(We suppose that infimum of an empty set is ∞). It should be mentioned that an analogous functional, defined using the Clarke subdifferentials, was introduced in [20] and then used in [10, 17].

A point $x \in X$ is a critical point of f if $0 \in \partial f(x)$.

A real number $c \in R^1$ is called a critical value of f if there exists a critical point x of f such that $f(x) = c$.

For each function $h : X \to R^1$ and each nonempty set $A \subset X$ set

$$\inf(h) = \inf\{h(z) : \ z \in X\}, \ \inf(h; A) = \inf\{h(z) : \ z \in A\}.$$

For each $x \in X$ and each $B \subset X$ put

$$d(x, B) = \inf\{||x - y|| : \ y \in B\}.$$

Let $f : X \to R^1$ be a function which is Lipschtzian on all bounded subsets of X and which satisfies the growth condition

$$\lim_{||x||\to\infty} f(x) = \infty. \tag{2.5}$$

Clearly, the function f is bounded from below.

Let $g : X \to R^1$ be a locally Lipschitzian function.

We say that the function g satisfies the Palais-Smale (P-S) condition on a set $M \subset X$ if for each normed-bounded sequence $\{z_i\}_{i=1}^{\infty} \subset M$ such that the sequence $\{g(z_i)\}_{i=1}^{\infty}$ is bounded and $\liminf_{i\to\infty} \Xi_g(z_i) = 0$ there exists a norm convergent subsequence of $\{z_i\}_{i=1}^{\infty}$ [1-3, 15, 20].

Let $c \in R^1$ be such that $g^{-1}(c) \neq \emptyset$.

We consider the following constrained minimization problems

$$(P_e) \qquad f(x) \to \min \text{ subject to } x \in g^{-1}(c)$$

and

$$(P_i) \qquad f(x) \to \min \text{ subject to } x \in g^{-1}((-\infty, c]).$$

We associate with these two problems the corresponding families of unconstrained minimization problems

$$(P_{\lambda e}) \qquad f(x) + \lambda|g(x) - c| \to \min, \ x \in X$$

and

$$(P_{\lambda i}) \qquad f(x) + \lambda \max\{g(x) - c, 0\} \to \min, \ x \in X$$

where $\lambda > 0$.

The following two theorems are our main results.

Theorem 2.1. *Assume that there exists $\gamma_* > 0$ such that the functions g and $-g$ satisfy the (P-S) condition on the set $g^{-1}([c-\gamma_*, c+\gamma_*])$ and the following property holds:*

If $x \in g^{-1}(c)$ is a critical point of the function g or a critical point of the function $-g$, then $f(x) > \inf(f; g^{-1}(c))$.

Then there exists a positive number $\bar{\lambda}$ such that for each $\epsilon > 0$ there exists $\delta \in (0, \epsilon)$ such that the following assertion holds.

If $\lambda > \bar{\lambda}$ and if $x \in X$ satisfies

$$f(x) + \lambda|g(x) - c| \leq \inf\{f(z) + \lambda|g(z) - c| : \ z \in X\} + \delta,$$

then there is $y \in g^{-1}(c)$ such that

$$||y - x|| \leq \epsilon \text{ and } f(y) \leq \inf(f; g^{-1}(c)) + \delta.$$

Theorem 2.2. *Assume that there exists $\gamma_* > 0$ such that the function g satisfies the (P-S) condition on the set $g^{-1}([c, c + \gamma_*])$ and the following property holds:*

If $x \in g^{-1}(c)$ is a critical point of the function g, then $f(x) > \inf(f; g^{-1}(-\infty, c])$.

Then there exists a positive number $\bar{\lambda}$ such that for each $\epsilon > 0$ there exists $\delta \in (0, \epsilon)$ such that the following assertion holds.

If $\lambda > \bar{\lambda}$ and if $x \in X$ satisfies

$$f(x) + \lambda \max\{g(x) - c, 0\} \leq \inf\{f(z) + \lambda \max\{g(z) - c, 0\} : \ z \in X\} + \delta,$$

then there is $y \in g^{-1}(-\infty, c])$ such that

$$||y - x|| \leq \epsilon \text{ and } f(y) \leq \inf(f; g^{-1}(-\infty, c]) + \delta.$$

3. Proofs of Theorems 2.1 and 2.2

We prove Theorems 2.1 and 2.2 simultaneously. Set

$$(3.1) \qquad A = g^{-1}(c) \text{ in the case of Theorem 2.1}$$

and

$$(3.2) \qquad A = g^{-1}((-\infty, c]) \text{ in the case of Theorem 2.2.}$$

For each $\lambda > 0$ define a function $\psi_\lambda : X \to R^1$ as

$$\psi_\lambda(z) = f(z) + \lambda|g(z) - c|, \; z \in X \tag{3.3}$$

in the case of Theorem 2.1 and as

$$\psi_\lambda(z) = f(z) + \lambda \max\{g(z) - c, 0\}, \; z \in X \tag{3.4}$$

in the case of Theorem 2.2.

Clearly, the function ψ_λ is locally Lipschitzian for all $\lambda > 0$. We show that there exists $\bar{\lambda} > 0$ such that the following property holds:

(P) For each $\epsilon \in (0, 1)$ there exists $\delta \in (0, \epsilon)$ such that for each $\lambda > \bar{\lambda}$ and each $x \in X$ which satisfies

$$\psi_\lambda(x) \le \inf(\psi_\lambda) + \delta$$

the set

$$\{y \in A : \; ||x - y|| \le \epsilon \text{ and } \psi_\lambda(y) \le \psi_\lambda(x)\}$$

is nonempty.

It is not difficult to see that the existence of $\bar{\lambda} > 0$ for which the property (P) holds implies the validity of Theorems 2.1 and 2.2.

Let us assume the contrary. Then for each natural number k there exist

$$\epsilon_k \in (0, 1), \; \lambda_k > k, \; x_k \in X \tag{3.5}$$

such that

$$\psi_{\lambda_k}(x_k) \le \inf(\psi_{\lambda_k}) + 2^{-1}\epsilon_k k^{-2} \tag{3.6}$$

and

$$\{z \in A : \; ||z - x_k|| \le \epsilon_k \text{ and } \psi_{\lambda_k}(y_k) \le \psi_{\lambda_k}(x_k)\} = \emptyset. \tag{3.7}$$

Let k be a natural number. It follows from (3.6) and Ekeland's variational principle [9] that there exists $y_k \in X$ such that

$$\psi_{\lambda_k}(y_k) \le \psi_{\lambda_k}(x_k), \tag{3.8}$$

$$||y_k - x_k|| \le (2k)^{-1}\epsilon_k, \tag{3.9}$$

$$\psi_{\lambda_k}(y_k) \le \psi_{\lambda_k)}(z) + k^{-1}||z - y_k|| \text{ for all } z \in X. \tag{3.10}$$

By (3.8), (3.9) and (3.7),

$$y_k \notin A \text{ for all natural numbers } k. \tag{3.11}$$

In the case of Theorem 2.2 we obtain that

$$g(y_k) > c \text{ for all natural numbers } k. \tag{3.12}$$

In the case of Theorem 2.1 we obtain that for each natural number k either $g(y_k) > c$ or $g(y_k) < c$.

In the case of Theorem 2.1 by extracting a subsequence and re-indexing we may assume that either $g(y_k) > c$ for all natural numbers k or $g(y_k) < c$ for all natural natural numbers k. Replacing g with $-g$ and c with $-c$ if necessary we may assume without loss of generality that (3.12) holds in the case of Theorem 2.1 too. Now (3.12) is valid in both cases.

Let k be a natural number. Then by (3.12) there is an open neighborhood V_k of y_k in X such that

$$g(z) > c \text{ for all } z \in V_k.$$

Together with (3.3), (3.4) and (3.10) this implies that for all $z \in V_k$

$$f(y_k) + \lambda_k(g(y_k) - c) = \psi_{\lambda_k}(y_k) \le f(z) + \lambda_k(g(z) - c) + k^{-1}||z - y_k||. \tag{3.13}$$

Put

$$\phi_k(z) = f(z) + \lambda_k g(z) + k^{-1}||z - y_k||, \; z \in V_k. \tag{3.14}$$

By (2.3), (2.2), (3.14) and (3.13),

$$0 \in \partial(\phi_k)(y_k). \tag{3.15}$$

It follows from (3.15), (3.14) and Theorem 3.36 of [13] that

$$0 \in \partial f(y_k) + \lambda_k \partial g(y_k) + k^{-1}\partial(||\cdot - y_k||)(y_k). \tag{3.16}$$

In view of (3.16) and Corollary 1.8.1 of [13],

$$\begin{aligned} 0 &\in \partial g(y_k) + \lambda_k^{-1}\partial f(y_k) + \lambda_k^{-1}k^{-1}\partial(||\cdot - y_k||)(y_k) \\ &\subset \partial g(y_k) + \lambda_k^{-1}\partial f(y_k) + \lambda_k^{-1}k^{-1}B_*(0,1). \end{aligned} \tag{3.17}$$

It follows from (3.3)-(3.6) and (3.8) that for all natural numbers k

$$f(y_k) \le \psi_{\lambda_k}(y_k) \le \inf(\psi_{\lambda_k}) + 1 \le \inf(\psi_{\lambda_k}; A) + 1 = \inf(f; A) + 1. \tag{3.18}$$

In view of this inequality and the growth condition (2.5) the sequence $\{y_k\}_{k=1}^{\infty}$ is bounded. Since the function f is Lipschitzian on bounded subsets of X it follows from the boundedness of the sequence $\{y_k\}_{k=0}^{\infty}$ and Corollary 1.8.1 of [13] that there exists $L > 0$ such that

$$\partial f(y_k) \subset B_*(0, L) \text{ for all natural numbers } k. \tag{3.19}$$

By (3.5), (3.17) and (3.19) for all natural numbers k,

$$0 \in \partial g(y_k) + \lambda_k^{-1}B_*(0, L) + k^{-1}B_*(0, 1)$$

and in view of (2.4)

$$\lim_{k\to\infty} \Xi_g(y_k) = 0. \tag{3.20}$$

By (3.3)-(3.6), (3.8) and (3.12) for all integers $k \ge 1$,

$$\begin{aligned} \inf(f) + \lambda_k(g(y_k) - c) &\le f(y_k) + \lambda_k(g(y_k) - c) = \psi_{\lambda_k}(y_k) \\ &\le \inf(f; A) + 1 \end{aligned}$$

and

$$0 < g(y_k) - c \le \lambda_k^{-1}[\inf(f; A) + 1 - \inf(f)] \to 0 \text{ as } k \to \infty. \tag{3.21}$$

Hence there is a natural number k_0 such that for all integers $k \ge k_0$

$$g(y_k) \in (c, c + \gamma_*]. \tag{3.22}$$

By (3.22), the boundedness of the sequence $\{y_k\}_{k=1}^{\infty}$ in the norm topology and the (P-S) condition there exists a strictly increasing sequence of natural numbers

$\{k_j\}_{j=1}^{\infty}$ such that $\{y_{k_j}\}_{j=1}^{\infty}$ converges in the norm topology to $y_* \in X$. In view of (3.21),

$$g(y_*) = c. \tag{3.23}$$

By (3.3), (3.4), (3.8) and (3.16),

$$f(y_*) = \lim_{j\to\infty} f(y_{k_j}) \leq \limsup_{j\to\infty} \psi_{\lambda_{k_j}}(y_{k_j})$$

$$\leq \limsup_{j\to\infty} \inf(\psi_{\lambda_{k_j}}) \leq \limsup_{j\to\infty} \inf(\psi_{\lambda_{k_j}}; A) = \inf(f; A).$$

Together with (3.23), (3.1) and (3.2) this implies that

$$f(y_*) = \inf(f; A). \tag{3.24}$$

We have already mentioned that the sequence $\{y_k\}_{k=1}^{\infty}$ is bounded. Choose a number $M > 4$ such that

$$\{y_k\}_{k=1}^{\infty} \subset B(0, M-2). \tag{3.25}$$

Since the function f is Lipschitz on bounded subsets of X there exists $L_0 \geq 1$ such that

$$|f(z_1) - f(z_2)| \leq L_0||z_1 - z_2|| \text{ for all } z_1, z_2 \in B(0, M). \tag{3.26}$$

Let k be a natural number. We may assume that

$$V_k \subset B(y_k, 1) \subset B(0, M). \tag{3.27}$$

By (3.13), (3.27), (3.26) and (3.15) for all $z \in V_k \setminus \{y_k\}$,

$$(g(z) - g(y_k))||z - y_k||^{-1} \geq ||z - y_k||^{-1}[f(y_k) - f(z)]\lambda_k^{-1} - k^{-1}\lambda_k^{-1}$$

$$\geq -L_0\lambda_k^{-1} - k^{-1} \geq -k^{-1}(1 + L_0).$$

By the relation above and the definition (2.2),

$$0 \in \widehat{\partial}_{a\gamma_k} g_k(y_k)$$

with

$$\gamma_k = k^{-1}(1 + L_0).$$

Together with (2.3) and the equality $y_* = \lim_{j\to\infty} y_{k_j}$ in the norm topology, $0 \in \partial g(y_*)$. This contradicts to the relation (3.23) and (3.24). The contradiction we have reached proves the existence of $\bar{\lambda} > 0$ for which the property (P) holds.

This completes the proofs of Theorems 2.1 and 2.2.

Acknowledgement

The author thanks B. S. Mordukhovich for useful discussions.

References

[1] S. Aizicovici, S. Reich and A. J. Zaslavski, *Convergence theorems for continuous descent methods*, Journal of Evolution Equaions **4** (2004), 139–156.

[2] T. Q. Bao and B. S. Mordukhovich, *Variational principles for set-valued mappings with applications to multiobjective optimization*, Control and Cybernetics **36** (2007), 531–562.

[3] T. Q. Bao and B. S. Mordukhovich, *Relative Pareto minimizers for multiobjective problems: existence and optimality conditions*, Math. Prog. A, published on-line.

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

[5] J. V.Burke, *Calmness and exact penalization*, SIAM J. Control Optim. **29** (1991), 493–497.

[6] J. V.Burke, *An exact penalization viewpoint of constrained optimization*, SIAM J. Control Optim. **29** (1991), 968–998.

[7] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Willey Interscience, New York, 1983.

[8] G. Di Pillo and L. Grippo, *Exact penalty functions in constrained optimization*, SIAM J. Control Optim. **27** (1989), 1333–1360.

[9] I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.

[10] I. I. Eremin, *The penalty method in convex programming*, Soviet Math. Dokl. **8** (1966), 459–462.

[11] M. Khaladi and J.-P. Penot, *Estimates of the exact penalty coefficient threshold*, Utilitas Mathematica **42** (1992), 147–161.

[12] B. S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech. **40** (1976), 960–969.

[13] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory*, Springer, Berlin, 2006.

[14] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, II: Applications*, Springer, Berlin, 2006.

[15] R. Palais *Lusternik-Schnirelman theory of Banach manifolds*, Topology, **5** (1966), 115–132.

[16] S. Reich and A. J. Zaslavski, *Porosity of the set of divergent descent methods*, Nonlinear Analysis **47** (2001), 3247–3258.

[17] S. Reich and A. J. Zaslavski, *Regular vector fields in Banach spaces*, Taiwanese J. Math. **12** (2008), 1165–1176.

[18] D. E. Ward, *Exact penalties and sufficient conditions for optimality in nonsmooth optimization*, J. Optim. Theory Appl. **57** (1988), 485–499.

[19] W. I. Zangwill, *Nonlinear programming via penalty functions*, Management Sci. **13** (1967), 344–358.

[20] A. J. Zaslavski, *On critical points of Lipschitz functions on smooth manifolds*, Siberian Math. J. **22** (1981), 63–68.

[21] A. J. Zaslavski, *A sufficient condition for exact penalty in constrained optimization*, SIAM Journal on Optimization **16** (2005), 250–262.

Manuscript received November 26, 2008
revised January 8, 2009

Alexander J. Zaslavski
Department of Mathematics, The Technion-Israel Institute of Technology, 32000 Haifa, Israel
E-mail address: ajzasl@tx.technion.ac.il