International Journal of Pure and Applied Mathematics

Volume 54 No. 2 2009, 153-160

EXISTENCE OF EXACT PENALTY FOR CONSTRAINED MINIMIZATION PROBLEMS ON BANACH SPACES

Alexander J. Zaslavski

Department of Mathematics
Technion-Israel Institute of Technology
Haifa, 32000, ISRAEL
e-mail: ajzasl@tx.technion.ac.il

Abstract: In this paper we use the penalty approach in order to study two constrained minimization problems. 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. In this paper we establish the existence of the exact penalty for an equality-constrained problem and an inequality-constrained problem in a Banach space.

AMS Subject Classification: 49M30, 90C26, 90C30
Key Words: Clarke's generalized gradient, Ekeland's variational principle, minimization problem, penalty function

1. Introduction

In this paper we use the penalty approach in order to study two constrained nonconvex minimization problems. The first problem is an equality-constrained problem in a Banach space and the second problem is an inequality-constrained problem in a Banach space. A penalty function is said to have the exact penalty property (see Boukari and Fiacco [1], Burke [2], Clarke [3], Di Pillo and Grippo [4], Luo et al [6] and Mordukhovich [7]) if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. In Zaslavski [10] assuming that an objective function f and a constraint function g are locally Lipschitz we established a very simple sufficient condition for the exact penalty property. In the present paper we

Received: May 12, 2006

extend the main results of Zaslavski [10] to the case when the functions f and g are Lipschitz only in a neighborhood of a solution of the constrained problem.

Let $(X, ||\cdot||)$ be a Banach space, $(X^*, ||\cdot||_*)$ its dual space and let $U \subset X$ be a nonempty open subset of X. Assume that a function $f : U \to R^1$ is Lipschitz. For each $x \in U$, let

$$f^0(x,h) = \limsup_{t\to 0^+, y\to x} [f(y+th) - f(y)]/t, \ h \in X$$

be the Clarke generalized directional derivative of f at the point x (see Clarke [3]), let

$$\partial f(x) = \{l \in X^* : \ f^0(x,h) \geq l(h) \text{ for all } h \in X\}$$

be Clarke's generalized gradient of f at x (see Clarke [3]) and set

$$\Xi_f(x) = \inf\{f^0(x,h) : \ h \in X \text{ and } ||h|| \leq 1\}$$

(see Zaslavski [9, 10]).

A point $x \in U$ is called a critical point of f if $0 \in \partial f(x)$. It is not difficult to see that $x \in U$ is a critical point of f if and only if $\Xi_f(x) = 0$.

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

It is known (see Clarke [3, Chapter 2, Section 2.3]) that $\partial(-f)(x) = -\partial f(x)$ for any $x \in U$. This equality implies that $x \in U$ is a critical point of f if and only if x is a critical point of $-f$ and $c \in R^1$ is a critical value of f if and only if $-c$ is a critical value of $-f$.

For each function $h : X \to R^1 \cup \{\infty\}$ and each nonempty $A \subset X$ put

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

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

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

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

$$B(x,r) = \{y \in X : \ ||x-y|| \leq r\}, \ B^0(x,r) = \{y \in X : \ ||y-x|| < r\},$$

$$B(r) = B(0,r), \ B^0(r) = B^0(0,r).$$

Consider a minimization problem $h(z) \to \min$, $z \in X$, where $h : X \to R^1$ is a continuous bounded from below function. If the space X is infinite-dimensional, 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 $h(x) \leq \inf(h) + \delta$.

Let $g : X \to R^1$ be a continuous function, $f : X \to R^1 \cup \{\infty\}$ be a lower semicontinuous function which is bounded from below and satisfies the following growth condition

$$\lim_{||x||\to\infty} f(x) = \infty \tag{1.1}$$

and let a real number c satisfy $g^{-1}(c) \neq \emptyset$.

We consider the constrained problems

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

and

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

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

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

and

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

where $\lambda > 0$.

We assume that there is $\gamma > 0$ such that the following assumption holds:

(A1) For each $r > 0$ the set $B(r) \cap g^{-1}([c-\gamma, c+\gamma])$ is compact.

In this paper we also use the following assumptions:

(A2) If $x \in g^{-1}(c)$ satisfies $f(x) = \inf(f; g^{-1}(c)) < \infty$, then there exists $r > 0$ such that the restrictions of the functions f, g to $B(x, r)$ are finite-valued and Lipschitz and x is not a critical point of g.

(A3) If $x \in g^{-1}(c)$ satisfies $f(x) = \inf(f; g^{-1}((-\infty, c])) < \infty$, then there exists $r > 0$ such that the restrictions of f, g to $B(x, r)$ are finite-valued and Lipschitz and x is not a critical point of g.

The next two theorems are the main results of the paper.

Theorem 1.1. *Assume that* $\inf(f; g^{-1}(c)) < \infty$ *and that (A1) and (A2) hold. Then there exists a positive number* Λ_0 *such that for each* $\epsilon > 0$ *there exists* $\delta \in (0, \epsilon)$ *such that the following assertion holds:*

If $\lambda \geq \Lambda_0$ *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 exists $y \in g^{-1}(c)$ *such that*

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

Theorem 1.2. *Assume that* $\inf(f; g^{-1}((-\infty, c])) < \infty$ *and that (A1) and (A3) hold. Then there exists a positive number* Λ_0 *such that for each* $\epsilon > 0$

there exists $\delta \in (0, \epsilon)$ *such that the following assertion holds:*

If $\lambda \geq \Lambda_0$ *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 exists $y \in g^{-1}((-\infty, c])$ *such that*

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

It should be mentioned that in the prototypes of Theorems 1.1 and 1.2 in Zaslavski [10] the assumption (A1) was replaced by a Palais-Smale (P-S) type condition (see Palais [8], Zaslavski [9, 10]) and instead of (A2) and (A3) it was assumed that c is not a critical value of g.

Theorems 1.1 and 1.2 will be proved in Section 2.

In this section we present several important results which easily follow from Theorems 1.1 and 1.2.

Theorems 1.1 and 1.2 imply the following result.

Theorem 1.3. *1. Assume that* $\inf(f; g^{-1}(c)) < \infty$ *and that (A1) and (A2) hold. Then there exists a positive number* Λ_0 *such that for each* $\lambda > \Lambda_0$ *and each sequence* $\{x_i\}_{i=1}^{\infty} \subset X$ *which satisfies*

$$\lim_{i \to \infty} [f(x_i) + \lambda |g(x_i) - c|] = \inf\{f(z) + \lambda |g(z) - c| :\ z \in X\}$$

there exists a sequence $\{y_i\}_{i=1}^{\infty} \subset g^{-1}(c)$ *such that*

$$\lim_{i \to \infty} f(y_i) = \inf(f; g^{-1}(c)) \text{ and } \lim_{i \to \infty} ||y_i - x_i|| = 0.$$

2. Assume that $\inf(f; g^{-1}((-\infty, c])) < \infty$ *and that (A1) and (A3) hold. Then there exists a positive number* Λ_0 *such that for each* $\lambda > \Lambda_0$ *and each sequence* $\{x_i\}_{i=1}^{\infty} \subset X$ *which satisfies*

$$\lim_{i \to \infty} [f(x_i) + \lambda \max\{g(x_i) - c, 0\}] = \inf\{f(z) + \lambda \max\{g(z) - c, 0\} :\ z \in X\}$$

there exists a sequence $\{y_i\}_{i=1}^{\infty} \subset g^{-1}((-\infty, c])$ *such that*

$$\lim_{i \to \infty} f(y_i) = \inf(f; g^{-1}((-\infty, c])) \text{ and } \lim_{i \to \infty} ||y_i - x_i|| = 0.$$

Assertion 1 of Theorem 1.3 implies the following result.

Theorem 1.4. *Assume that* $\inf(f; g^{-1}(c)) < \infty$*, (A1) and (A2) hold and that there exists* $\bar{x} \in g^{-1}(c)$ *for which the following conditions hold:*

$$f(\bar{x}) = \inf(f; g^{-1}(c));$$

any sequence $\{x_n\}_{n=1}^{\infty} \subset g^{-1}(c)$ *which satisfies* $\lim_{n \to \infty} f(x_n) = \inf(f; g^{-1}(c))$ *converges to* $\bar{x}$ *in the norm topology.*

Then there exists $\Lambda_0 > 0$ *such that for each* $\lambda > \Lambda_0$ *the point* $\bar{x}$ *is a unique solution of the minimization problem*

$$f(z) + \lambda|g(z) - c| \to \min, \ z \in X.$$

Assertion 2 of Theorem 1.3 implies the following result.

Theorem 1.5. *Assume that* $\inf(f; g^{-1}((-\infty, c])) < \infty$, *(A1) and (A3) hold and that there exists* $\bar{x} \in g^{-1}((-\infty, c])$ *for which the following conditions hold:*

$$f(\bar{x}) = \inf(f; g^{-1}((-\infty, c]));$$

any sequence $\{x_n\}_{n=1}^{\infty} \subset g^{-1}((-\infty, c])$ *which satisfies*

$$\lim_{n\to\infty} f(x_n) = \inf(f; g^{-1}((-\infty, c]))$$

converges to $\bar{x}$ *in the norm topology.*

Then there exists $\Lambda_0 > 0$ *such that for each* $\lambda > \Lambda_0$ *the point* $\bar{x}$ *is a unique solution of the minimization problem*

$$f(z) + \lambda \max\{g(z) - c, 0\} \to \min, \ z \in X.$$

2. Proofs of Theorems 1.1 and 1.2

We prove Theorems 1.1 and 1.2 simultaneously.

Since the function f is bounded from below there exists $\bar{a} > 0$ such that

$$f(x) \geq -\bar{a} \text{ for all } x \in X. \tag{2.1}$$

Set

$$A = g^{-1}(c) \text{ in the case of Theorem 1.1},$$

and

$$A = g^{-1}((-\infty, c]) \text{ in the case of Theorem 1.2.}$$

Clearly A is a nonempty closed subset of $(X, ||\cdot||)$. For each $\lambda > 0$ define a function $\psi_\lambda : X \to R^1 \cup \{\infty\}$ as follows:

$$\psi_\lambda(z) = f(z) + \lambda|g(z) - c|, \ z \in X, \tag{2.2}$$

in the case of Theorem 1.1, and

$$\psi_\lambda(z) = f(z) + \lambda \max\{g(z) - c, 0\}, \ z \in X, \tag{2.3}$$

in the case of Theorem 1.2.

Clearly, the function ψ_λ is lower semicontinuous for all $\lambda > 0$.

We show that there is $\Lambda_0 > 0$ such that the following property holds:

(P1) For each $\epsilon > 0$ there exists $\delta \in (0, \epsilon)$ such that for each $\lambda \geq \Lambda_0$ and each $x \in X$ which satisfies

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

there is $y \in A$ for which

$$||y - x|| \leq \epsilon, \ \psi_\lambda(y) \leq \inf(\psi_\lambda) + \epsilon.$$

It is not difficult to see that (P1) implies Theorems 1.1 and 1.2.

Let us assume that there is no $\Lambda_0 > 0$ for which (P1) holds. Then for each natural number k there exist

$$\epsilon_k \in (0, 1), \ \lambda_k \geq k \text{ and } x_k \in X \tag{2.4}$$

such that

$$\psi_{\lambda_k}(x_k) \leq \inf(\psi_{\lambda_k}) + 8^{-1}\epsilon_k k^{-2} \tag{2.5}$$

and

$$\{z \in A \cap B(x_k, \epsilon_k) : \psi_{\lambda_k}(z) \leq \inf(\psi_{\lambda_k}) + \epsilon_k\} = \emptyset. \tag{2.6}$$

Let k be a natural number. It follows from (2.5) and Ekeland's variational principle (see Ekeland [5]) that there is $y_k \in X$ such that

$$\psi_{\lambda_k}(y_k) \leq \psi_{\lambda_k}(x_k), \tag{2.7}$$

$$||y_k - x_k|| \leq 2^{-1}k^{-1}\epsilon_k, \tag{2.8}$$

$$\psi_{\lambda_k}(y_k) \leq \psi_{\lambda_k}(z) + k^{-1}||z - y_k|| \text{ for all } z \in X. \tag{2.9}$$

By (2.5)-(2.8),

$$y_k \notin A \text{ for all natural numbers } k. \tag{2.10}$$

In the case of Theorem 1.2 we obtain that

$$g(y_k) > c \text{ for all natural numbers } k. \tag{2.11}$$

In the case of Theorem 1.1 we obtain that for each natural number k

$$\text{either } g(y_k) > c \text{ or } g(y_k) < c.$$

In the case of Theorem 1.1 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 numbers k. Replacing g by $-g$ and c by $-c$, if necessary, we may assume without loss of generality that (2.11) is valid in the case of Theorem 1 too. Now (2.11) is valid in both cases.

It follows from (2.2), (2.3), (2.5), (2.7) and the definition of A that for each natural number k

$$f(y_k) \le \psi_{\lambda_k}(y_k) \le \inf(\psi_{\lambda_k}) + 1$$
$$\le \inf\{\psi_{\lambda_k}(z) :\ z \in A\} + 1 = \inf\{f(z) :\ z \in A\} + 1. \quad (2.12)$$

By (1.1) and (2.12) there exists $r_0 > 0$ such that

$$||y_k|| \le r_0 \text{ for all integers } k \ge 1. \quad (2.13)$$

In view of (2.1)-(2.3), (2.11) and (2.12) for all natural numbers k

$$-\bar{a} + \lambda_k(g(y_k) - c) \le f(y_k) + \lambda_k(g(y_k) - c) = \psi_{\lambda_k}(y_k) \le \inf(f; A) + 1,$$
$$0 < g(y_k) - c < \lambda_k^{-1}(\inf(f; A) + \bar{a} + 1). \quad (2.14)$$

It follows from (2.13), (2.14) and (A1) that the sequence $\{y_k\}_{k=1}^{\infty}$ possesses a norm convergent subsequence. Now extracting a subsequence and re-indexing we may assume without loss of generality that there exists

$$y_* = \lim_{k\to\infty} y_k \quad (2.15)$$

in the norm topology. Relations (2.14) and (2.15) imply that

$$g(y_*) = \lim_{k\to\infty} g(y_k) = c \text{ and } y_* \in A. \quad (2.16)$$

By the lower semicontinuity of f, (2.2), (2.3), (2.5), (2.7) and (2.15)

$$f(y_*) \le \liminf_{k\to\infty} f(y_k) \le \liminf_{k\to\infty}(\psi_{\lambda_k}) \le \liminf_{k\to\infty}(\psi_{\lambda_k}; A) = \inf(f; A). \quad (2.17)$$

Together with (2.16) this implies that

$$f(y_*) = \inf(f; A). \quad (2.18)$$

In view of (2.18), (2.16), (A2) and (A3) there exist $r_1 > 0$, $L_1 > 1$ such that

$$|f(z_1) - f(z_2)|,\ |g(z_1) - g(z_2)| \le L_1||z_1 - z_2|| \text{ for all } z_1, z_2 \in B(y_*, r_1). \quad (2.19)$$

By (2.15) there exists a natural number k_1 such that

$$||y_k - y_*|| \le r_1/4 \text{ for all integers } k \ge k_1. \quad (2.20)$$

Let $k \ge k_1$ be an integer. By (2.11) there is an open neighborhood V_k of y_k in X with the norm topology such that

$$V_k \subset B(y_k, r_1/4),\ g(z) > c \text{ for all } z \in V_k. \quad (2.21)$$

Relations (2.2), (2.3) and (2.21) imply that for all $z \in V_k$

$$\psi_{\lambda_k}(z) = f(z) + \lambda_k(g(z) - c). \quad (2.22)$$

It follows from the choice of r_1, L_1, (2.9), (2.19) and (2.22) that

$$0 \in \partial f(y_k) + \lambda_k \partial g(y_k) + k^{-1}\{l \in X^* :\ ||l||_* \le 1\}. \quad (2.23)$$

By (2.4), (2.19) and (2.23)

$$0 \in \lambda_k^{-1}\partial f(y_k) + \partial g(y_k) + \{l \in X^* :\ ||l||_* \le k^{-1}\}$$
$$\subset \partial g(y_k) + \lambda_k^{-1}\{l \in X_* :\ ||l||_* \le L_1\} + \{l \in X^* :\ ||l|| \le k^{-1}\}$$

and

$$\liminf_{k\to\infty} \Xi_g(y_k) \geq 0.$$

Combined with (2.15) and the definition of Ξ_g this implies that

$$\Xi_g(y_*) = 0.$$

Together with (2.16) and (2.18) this equality contradicts (A2) and (A3). The contradiction we have reached proves that there is $\Lambda_0 > 0$ such that (P1) holds. This completes the proof of Theorems 1.1 and 1.2.

References

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

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

[3] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Willey Interscience (1983).

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

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

[6] Z.Q. Luo, J.-S. Pang, D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press (1996).

[7] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin (2006).

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

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

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