TAIWANESE JOURNAL OF MATHEMATICS
Vol. 15, No. 5, pp. 2235-2243, October 2011
This paper is available online at http://tjm.math.ntu.edu.tw/index.php/TJM

LOCAL SOLUTIONS OF CONSTRAINED MINIMIZATION PROBLEMS AND CRITICAL POINTS OF LIPSCHITZ FUNCTIONS

Alexander J. Zaslavski

Abstract. In this paper we use the penalty approach in order to study two constrained nonconvex minimization problems with locally Lipschitzian objective and constraint functions in a Banach space. We show that a local minimizer of the constrained minimization problem which is not a critical point of the constraint function is also a local minimizer of a corresponding unconstrained penalized problem if a penalty coefficient is large enough.

1. Introduction and the Main Result

Penalty methods are an important and useful tool in constrained optimization. See, for example, [1, 2, 4] and the references mentioned there. In this paper we use the penalty approach in order to study two constrained nonconvex minimization problems with locally Lipschitzian objective and constraint functions in a Banach space. The first problem is an equality-constrained problem and the second one is an inequality-constrained problem. 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 [5] and Zangwill [12] for use in the development of algorithms for nonlinear constrained optimization. Since that time exact penalty functions have continued to play a key role in the theory of mathematical programming [1-4, 6-10]. A local exactness of penalties was studied in [6, 8, 10]. For more discussions and various applications of exact penalization to various constrained optimization problems see [1, 2, 4, 10].

Usually the exact penalty property is related to calmness of the perturbed constraint function. In [14] we use an assumption of different nature which is not difficult to verify. In particular, we show in [14] that the problem $f(x) \to \min$

Received April 30, 2010, accepted June 15, 2010.
Communicated by Boris S. Mordukhovich.
2010 *Mathematics Subject Classification*: 49M30, 90C26, 90C30.
Key words and phrases: Clarke's generalized gradient, Critical point, Lipschitz function, Minimization problem, Penalty function.

subject to $g(x) = c$ possesses the exact penalty if the real number c is not a critical value of the function g. In other words the set $g^{-1}(c)$ does not contain a critical point of the function g. Note that in [14] we used the notion of a critical point of a Lipschitz function introduced in [13]. The result of [14] was generalized in [15] for a constrained minimization problem with an arbitrary number of mixed constraints. Moreover, in [15] we do not assume that the set $g^{-1}(c)$ does not contain a critical point of the function g. Instead of it we suppose that the set $g^{-1}(c)$ does not contain a critical point of the function g which is a minimizer of the constrained minimization problem. In the present paper we make another development of the result of [14] and establish the existence of a local penalty. More precisely, we show that a local minimizer of the constrained minimization problem which is not a critical point of the constraint function is also a local minimizer of a corresponding unconstrained penalized problem if a penalty coefficient is large enough.

Let $\mathbf{R}$ be the set of all real numbers, $(X, ||\cdot||)$ be a Banach space, $(X^*, ||\cdot||_*)$ its dual space and let $\phi : X \to \mathbf{R}$ be a locally Lipschitzian function. For each $x \in X$ let

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

be the Clarke generalized directional derivative of ϕ at the point x [3], let

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

be Clarke's generalized gradient of ϕ at x [3] and set

$$\Xi_\phi(x) = \inf\{\phi^0(x,h) : \ h \in X \text{ and } ||h|| \leq 1\} \tag{1.1}$$

(see [13, 14]).

A point $x \in X$ is called a critical point of ϕ if $0 \in \partial\phi(x)$ [13, 14]. It is not difficult to see that $x \in X$ is a critical point of ϕ if and only if $\Xi_\phi(x) \geq 0$.

A real number $c \in \mathbf{R}$ is called a critical value of ϕ if there is a critical point x of ϕ such that $\phi(x) = c$.

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

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

Suppose that a function $f : X \to \mathbf{R} \cup \{\infty\}$, a function $g : X \to \mathbf{R}$ is locally Lipschitzian and that a real number c is such that the set $g^{-1}(c)$ is nonempty.

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

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

and

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

where $\lambda > 0$.

The following theorem is our main result.

Theorem 1.1. *Assume that*

$$(1.2) \qquad \bar{x} \in X \textit{ satisfies } g(\bar{x}) = c,$$

$\bar{x}$ is not a critical point of g and that there exists $\bar{r} > 0$ such that the following properties hold: the function f is finite-valued and Lipschitzian on the set $B(\bar{x}, \bar{r})$;

$$(1.3) \qquad f(x) \geq f(\bar{x}) \textit{ for each } x \in B(\bar{x}, \bar{r}) \cap g^{-1}(c).$$

Then there exist $r_1 > 0$ and $\Lambda_0 > 0$ such that if $\lambda \geq \Lambda_0$ and if $x \in B(\bar{x}, r_1)$ satisfies

$$f(x) + \lambda|g(x) - c| \leq f(\bar{x}),$$

then $g(x) = c$.

Corollary 1.1. *Assume that all the assumptions of Theorem 1.1 hold and let $r_1 > 0$ and $\Lambda_0 > 0$ be as guaranteed by Theorem 1.1 Then if $\lambda \geq \Lambda_0$ and if $x \in B(\bar{x}, r_1)$ satisfies*

$$f(x) + \lambda \max\{g(x) - c, 0\} \leq f(\bar{x}),$$

then $g(x) \leq c$. Denote by $\mathrm{co}(A)$ the convex hull of a set $A \subset X^*$.

Theorem 1.1 implies the following result which establishes a necessary optimality condition for the problem (P_e).

Proposition 1.1. *Assume that all the assumptions of Theorem 1.1 hold. Then there is $\Lambda_0 > 0$ such that*

$$0 \in \partial f(\bar{x}) + \Lambda_0 co(\partial g(\bar{x}) \cup (-\partial g(\bar{x})).$$

Proof. Let $r_1 > 0$ and $\Lambda_0 > 0$ be as guaranteed by Theorem 1.1. We may assume that $r_1 < \bar{r}$. Then for each $x \in B(\bar{x}, r_1)$

$$f(\bar{x}) + \Lambda_0|g(\bar{x}) - c| = f(\bar{x}) \leq f(x) + \Lambda_0|g(x) - c|.$$

This implies that

$$0 \in \partial f(\bar{x}) + \Lambda_0 \partial(|g(\cdot) - c|)(\bar{x}) \subset \partial f(\bar{x}) + \Lambda_0 \text{co}(\partial g(\bar{x}) \cup (-\partial g(\bar{x})).$$

Proposition 1.1 is proved. ∎

Corollary 1.1 implies the following result which establishes a necessary optimality condition for the problem (P_i).

Proposition 1.2. *Assume that all the assumptions of Theorem 1.1 hold and that*

$$f(x) \geq f(\bar{x}) \text{ for each } x \in B(\bar{x}, \bar{r}) \cap g^{-1}((-\infty, c]).$$

Then there is $\Lambda_0 > 0$ *such that*

$$0 \in \partial f(\bar{x}) + \Lambda_0 \text{co}(\partial g(\bar{x}) \cup \{0\}).$$

Proof. Let $r_1 > 0$ and $\Lambda_0 > 0$ be as guaranteed by Theorem 1.1. We may assume that $r_1 < \bar{r}$. Corollary 1.1 implies that for each $x \in B(\bar{x}, r_1)$

$$f(\bar{x}) + \Lambda_0 \max\{g(\bar{x}) - c, 0\} = f(\bar{x}) \leq f(x) + \Lambda_0 \max\{g(x) - c, 0\}.$$

This implies that

$$0 \in \partial f(\bar{x}) + \Lambda_0 \partial(\max\{g(\cdot) - c, 0\})(\bar{x}) \subset \partial f(\bar{x}) + \Lambda_0 \text{co}(\partial g(\bar{x}) \cup \{0\}).$$

Proposition 1.2 is proved. ∎

2. An Auxiliary Result

Let $(Y, ||\cdot||)$ and $(Z, ||\cdot||)$ be Banach spaces, $A \subset Y$ and $B \subset Z$. We say that $h : A \to B$ ia an $\mathcal{L}$-mapping if for each $x \in A$ there exists $r > 0$ such that the restriction $h : A \cap B(x, r) \to B$ is Lipschitz.

Assume that $g : X \to \mathbf{R}$ is a locally Lipschitz function. In the sequel we use the following auxiliary result obtained in [16].

Lemma 2.1. *Assume that* $x_0 \in X$, $\delta > 0$ *and that* $\Xi_g(x_0) < -\delta$. *Then there exist* $r > 0$ *and an* $\mathcal{L}-$*mapping* $V : X \to X$ *such that*

$$||V(x)|| \leq 2 \text{ for all } x \in X,$$

$$g^0(x, V(x)) \leq 0 \text{ for all } x \in X,$$

$$g^0(x, V(x)) \leq -\delta \text{ for all } x \in B(x_0, r).$$

3. Proof of Theorem 1.1

There exists $M_0 > 0$ such that

$$|f(z)| \le M_0 \text{ for all } z \in B(\bar{x}, \bar{r}) \tag{3.1}$$

and

$$|f(z_1) - f(z_2)| \le M_0||z_1 - z_2|| \text{ for all } z_1, z_2 \in B(\bar{x}, \bar{r}). \tag{3.2}$$

Since $\bar{x}$ is not a critical point of g there is $\delta > 0$ such that

$$\Xi_g(\bar{x}) < -\delta. \tag{3.3}$$

By Lemma 2.1 and (3.3) there exist $r_0 > 0$ and an $\mathcal{L}$-mapping $V : X \to X$ such that

$$||V(x)|| \le 2 \text{ for all } x \in X, \tag{3.4}$$

$$g^0(x, V(x)) \le 0 \text{ for all } x \in X \tag{3.5}$$

and

$$g^0(x, V(x)) \le -\delta \text{ for all } x \in B(\bar{x}, r_0). \tag{3.6}$$

We may assume without loss of generality that

$$r_0 < \bar{r}. \tag{3.7}$$

It was shown in [11] that the mapping V generates a flow $\sigma : \mathbf{R} \times X \to X$ such that the mapping σ is continuous and that

$$(d/dt)\sigma(t, x) = V(\sigma(t, x)) \text{ for all } x \in X \text{ and all } t \in \mathbf{R}. \tag{3.8}$$

Assume that

$$x \in X, \ t_1, t_2 \in \mathbf{R}, \text{ and } t_1 < t_2. \tag{3.9}$$

By the properties of the Clarke generalized directional derivative [3]

$$g(\sigma(t_2, x)) - g(\sigma(t_1, x)) \le (t_2 - t_1)l((d\sigma/dt)(\sigma(s, x))), \tag{3.10}$$

where

$$s \in [t_1, t_2] \text{ and } l \in \partial g(\sigma(s, x)). \tag{3.11}$$

By (3.10), (3.11), (3.8), (3.9) and (3.5),

$$g(\sigma(t_2,x)) - g(\sigma(t_1,x)) = l(V(\sigma(s,x)))(t_2 - t_1) \le 0. \tag{3.12}$$

Thus

$$g(\sigma(t_2,x)) - g(\sigma(t_1,x)) \le 0$$

$$\text{for all } x \in X, \text{ each } t \in \mathbf{R} \text{ and each } t_2 > t_1. \tag{3.13}$$

There are

$$\tau_1 \in (0,1),\ r_1 \in (0, r_0) \tag{3.14}$$

such that

$$\begin{aligned} &||\sigma(t,x) - \bar{x}|| < r_0 \\ &\text{for each } t \in [-\tau_1, \tau_1] \text{ and each } x \in B(\bar{x}, r_1). \end{aligned} \tag{3.15}$$

Assume now that

$$x \in B(\bar{x}, r_1),\ t_1, t_2 \in [-\tau_1, \tau_1] \text{ and } t_1 < t_2. \tag{3.16}$$

By the properties of the Clarke generalized directional derivative [3], (3.8) and (3.16),

$$g(\sigma(t_2,x)) - g(\sigma(t_1,x)) = l(d\sigma/dt(\sigma(s,x)))(t_2 - t_1), \tag{3.17}$$

where

$$s \in [t_1, t_2] \text{ and } l \in \partial g(\sigma(s,x)). \tag{3.18}$$

By (3.17) and (3.8),

$$g(\sigma(t_2,x)) - g(\sigma(t_1,x)) = l(V(\sigma(s,x)))(t_2 - t_1). \tag{3.19}$$

In view of (3.18), (3.16) and (3.15),

$$\sigma(s,x) \in B(\bar{x}, r_0). \tag{3.20}$$

By (3.20) and (3.6),

$$g^0(\sigma(s,x), V(\sigma(s,x))) \le -\delta. \tag{3.21}$$

It follows from (3.19), (3.18) and (3.21) that

$$g(\sigma(t_2,x)) - g(\sigma(t_1,x)) \le -\delta(t_2 - t_1) \tag{3.22}$$

for each $x \in B(\bar{x}, r_1)$ and each $t_1, t_2 \in [-\tau_1, \tau_2]$ satisfying $t_1 < t_2$. By (3.22) for each $x \in B(\bar{x}, r_1)$,

$$g(\sigma(\tau_1, x)) \le g(x) - \delta\tau_1$$

and

$$g(\sigma(-\tau_1, x)) \geq g(x) + \delta\tau_1. \tag{3.23}$$

Choose a positive number Λ_0 such that

$$\Lambda_0 > 2M_0\delta^{-1}\tau_1^{-1}. \tag{3.24}$$

Assume that

$$\lambda \geq \Lambda_0,\ x \in B(\bar{x}, r_1) \tag{3.25}$$

and

$$f(x) + \lambda|g(x) - c| \leq f(\bar{x}). \tag{3.26}$$

We show that $g(x) = c$. Assume the contrary. Then

$$g(x) \neq c. \tag{3.27}$$

By (3.25), (3.26), (3.14), (3.7) and (3.1),

$$\Lambda_0|g(x) - c| \leq f(\bar{x}) + M_0 \leq 2M_0$$

and

$$|g(x) - c| \leq 2M_0\Lambda_0^{-1}. \tag{3.28}$$

In view of (3.23), (3.25), (3.28) and (3.24),

$$g(\sigma(\tau_1, x)) \leq g(x) - \delta\tau_1 \leq c + 2M_0\Lambda_0^{-1} - \delta\tau_1 < c$$

and

$$g(\sigma(-\tau_1, x)) \geq \delta\tau_1 + g(x) \geq \delta\tau_1 + c - 2M_0\Lambda_0^{-1} > c.$$

It follows from the inequality above, (3.22) and (3.25) that there is a unique

$$s \in [-\tau_1, \tau_1] \tag{3.29}$$

such that

$$g(\sigma(s, x)) = c. \tag{3.30}$$

In view of (3.30), (3.29), (3.22) and (3.25),

$$|c - g(x)| = |g(\sigma(s, x)) - g(\sigma(0, x))| \geq \delta|s|. \tag{3.31}$$

By (3.8) and (3.5),

$$\begin{aligned} ||x - \sigma(s, x)|| &= ||\sigma(0, x) - \sigma(s, x)|| \\ &\leq |\int_0^s ||V(\sigma(t, x))||dt| \leq 2|s|. \end{aligned} \tag{3.32}$$

By (3.25), (3.29), (3.14) (3.15), (3.7), (3.2) and (3.32),

$$|f(x) - f(\sigma(s,x))| \leq M_0||x - \sigma(s,x)|| \leq 2M_0|s|. \tag{3.33}$$

It follows from (3.30), (3.33), (3.31), (3.27) and (3.14) that

$$\begin{aligned} &f(\sigma(s,x)) + \lambda|g(\sigma(s,x)) - c| \\ &= f(\sigma(s,x)) \leq f(x) + 2M_0|s| \\ &\leq f(x) + 2M_0\delta^{-1}|c - g(x)| < f(x) + \Lambda_0|c - g(x)|. \end{aligned}$$

Together with (1.3), (3.30), (3.29), (3.25), (3.15) and (3.7) this implies that

$$f(\bar{x}) \leq f(\sigma(s,x)) < f(x) + \lambda|g(x) - c|.$$

This contradicts (3.26). The contraction we have reached proves that $g(x) = c$. Theorem 1.1 is proved.

ACKNOWLEDGMENTS

The author thanks the referee for useful comments.

REFERENCES

1. D. Boukari and 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 and L. Grippo, Exact penalty functions in constrained optimization, *SIAM J. Control Optim.*, **27** (1989), 1333-1360.
5. I. I. Eremin, The penalty method in convex programming, *Soviet Math. Dokl.*, **8** (1966), 459-462.
6. S.-P. Han and O. L. Mangasarian, Exact penalty function in nonlinear programming, *Math. Programming*, **17** (1979), 251-269.
7. J.-B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer, Berlin, 1993,
8. A. D. Ioffe, Necessary and sufficient conditions for a local minimum. I. A reduction theorem and first order conditions, *SIAM J. Control Optim.*, **17** (1979), 245-250.
9. B. S. Mordukhovich, Penalty functions and necessary conditions for the extremum in nonsmooth and nonconvex optimization problems, *Uspekhi Math. Nauk*, **36** (1981), 215-216.

10. B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin, 2006.

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

12. W. I. Zangwill, Nonlinear programming via penalty functions, *Management Sci.*, **13** (1967), 344-358.

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

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

15. A. J. Zaslavski, Existence of exact penalty for optimization problems with mixed constraints in Banach spaces, *J. Math. Anal. Appl*, **324** (2006), 669-681.

16. A. J. Zaslavski, Stability of exact penalty for classes of constrained minimization problems in Banach spaces, *Taiwanese J. Math.*, **12** (2008), 1493-1510.

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