Journal of Global Optimization (2006) 36:637–652
DOI 10.1007/s10898-004-1937-y

Convergence Analysis of a Class of Penalty Methods for Vector Optimization Problems with Cone Constraints*

X.X. HUANG[1], X.Q. YANG[2] and K.L. TEO[2]
[1]*School of Management, Fudan University, Shanghai 200433, China*
[2]*Department of Applied Mathematics, HongKong PolytechnicUniversity, Kowloon, HongKong*

(Received 11 November 2003; accepted in revised form 28 July 2004)

Abstract. In this paper, we consider convergence properties of a class of penalization methods for a general vector optimization problem with cone constraints in infinite dimensional spaces. Under certain assumptions, we show that any efficient point of the cone constrained vector optimization problem can be approached by a sequence of efficient points of the penalty problems. We also show, on the other hand, that any limit point of a sequence of approximate efficient solutions to the penalty problems is a weekly efficient solution of the original cone constrained vector optimization problem. Finally, when the constrained space is of finite dimension, we show that any limit point of a sequence of stationary points of the penalty problems is a KKT stationary point of the original cone constrained vector optimization problem if Mangasarian–Fromovitz constraint qualification holds at the limit point.

Key words: Convergence, Efficiency, Level-compactness, Penalty method, Vector optimization with cone constraints

1. Introduction and Preliminaries

It is well-known that penalty method is popular and effective for constrained optimization problems. Traditional penalty methods use l_2 and l_1 penalty functions (see, e.g., [7, 8]). Recent interest in lower-order penalty functions is motivated by the fact that lower-order penalty functions usually require weaker conditions to guarantee their exactness than the classical l_1 penalty function (see, e.g., [14, 16, 18, 21, 22]).

In scalar optimization, necessary and sufficient conditions were derived for various exact penalty functions, including l_1 exact penalization [2], nonlinear exact penalization [21] and lower order penalization [16]. These conditions were expressed in the form of calmness-types conditions of the constrained optimization problem.

*This work is supported by the Postdoctoral Fellowship of Hong Kong Polytechnic University.

In vector optimization in finite dimensional spaces, penalty methods were studied in [14, 20, 23]. Necessary and sufficient conditions were developed for the exactness of a class of nonlinear penalty functions for constrained vector optimization problems [14]. Most recently, for vector optimization problems with cone constraints in infinite dimensional spaces, we established an equivalence between (local) calmness property of the cone constrained optimization problem and (local) exact penalization property of a class of penalty functions. This calmness condition was further applied to derive KKT conditions for a local weakly efficient solution to the cone constrained vector optimization problem (see [12]).

Convergence analysis for a class of nonlinear penalty methods for constrained scalar optimization problems was carried out in [15, 24]. In [13] we studied the convergence property (via first-order necessary optimality conditions) of a class of nonlinear penalty methods for constrained multiobjective optimization problems.

In this paper, we will carry out convergence analysis for the class of penalty methods proposed in [12]. More specifically, under certain assumptions, we show that any efficient point of the cone constrained vector optimization problem can be approached by a sequence of efficient points of the penalty problems. We note that these assumptions are weaker than those used in [14] for strong duality results. We also show that any limit point of a sequence of approximate efficient solutions to the penalty problems is a weakly efficient solution of the original cone constrained vector optimization problem. Finally, under some conditions, we show that any limit point of a sequence of stationary points (points that satisfy first-order necessary conditions) of the penalty problems is a KKT stationary point of the original cone constrained vector optimization problem if the Mangasarian–Fromovitz constraint qualification holds at the limit point. It is worth nothing that the Mangasarian–Fromovitz condition is weaker than the linear independence condition of the gradients of active constraint functions used in [24] even when the constrained vector optimization problem considered in this paper reduces to the scalar constrained optimization problem considered in [24].

In the remainder of this section, we present the problem, some related concepts and notations.

Let X, Y and Z be normed spaces. Let Y^* and Z^* be the dual spaces of Y and Z, respectively. Let $C \subset Y$ be a nontrivial pointed, closed and convex cone with nonempty interior int$C$ and $e \in \text{int}C$ be fixed. Let $K \subset Z$ be a closed and convex cone with nonempty interior int$K$ and $k^0 \in \text{int}K$ be fixed. C and K induce orders $\leqslant_C$ and $\leqslant_K$ in Y and Z, respectively. That is, for any $y_i \in Y$, $z_i \in Z (i = 1, 2)$, $y_1 \leqslant_C y_2$ iff $y_2 - y_1 \in C$ and $z_1 \leqslant_K z_2$ iff $z_2 - z_1 \in K$. Denote by $C^* = \{\lambda \in Y^* : \lambda(y) \geqslant 0, \forall y \in C\}$ and $K^* = \{\mu \in Z^* : \mu(z) \geqslant 0, \forall z \in K\}$ the positive polar cones of C and K, respectively. Let $C^0 = \{\lambda \in C^* : \|\lambda\| = 1\}$.

DEFINITION 1.1. Let V be a normed space and $K_1 \subset V$ be a cone. We say that K_1 admits a compact base if there exists a compact set $K_2 \subset K_1$ such that $0 \notin K_2$ and $K_1 = \{\beta k : \beta \geqslant 0, k \in K_2\}$.

Let $Y_1 \subset Y$. $\bar{y} \in Y$ is called a infimum point of Y_1 if (i) there exists a sequence $\{y_n\} \subset Y_1$ such that $y_n \to \bar{y}$; and (ii) $y - \bar{y} \notin -C \backslash \{0\}$, $\forall y \in Y_1$. The set of all infimum points of Y_1 is denoted by $\inf Y_1$. If $\bar{y} \in (\inf Y_1) \cap Y_1$, then it is called an efficient point of Y_1.

Consider the following vector optimization problem with cone constraints

$$\begin{aligned} &(\mathrm{VP}) \ \inf f(X) \\ &\text{s.t. } x \in X_1, \\ &g(x) \in -K, \end{aligned}$$

when $X_1 \subset X$ is nonempty and closed, $f : X \to Y$, $g : X \to Z$ are vector-valued functions.

Throughout the paper, we assume that $\lambda(f(x))$ is lower semicontinuous (l.s.c. for short) on X_1 for any $\lambda \in C^*$ and $g(x)$ is continuous on X_1 unless stated otherwise.

Denote by X_0 the feasible set of (VP), i.e.,

$$X_0 = \{x \in X_1 : g(x) \in -K\}.$$

We assume throughout that $X_0 \neq \emptyset$. Denote by E' the set of infimum points of (VP), i.e., $E' = \inf f(X_0)$. Let E denote the set of efficient points of $f(X_0)$. E is also referred to as the set of efficient points of (VP). Clearly, $E \subset E'$.

A point $\bar{x} \in X_0$ is said to be a (weakly) efficient solution to (VP) if

$$f(x) - f(\bar{x}) \notin -C \backslash \{0\} (\text{resp.} - \mathrm{int} C) \quad \forall x \in X_0.$$

A point $\bar{x} \in X_0$ is said to be a local (weakly) efficient solution to (VP) if there exists a neighbourhood U of $\bar{x}$ such that

$$f(x) - f(\bar{x}) \notin -C \backslash \{0\} (\text{resp.} - \mathrm{int} C) \quad \forall x \in X_0 \cap U.$$

Clearly, a (local) efficient solution of (VP) is a (local) weakly efficient solution of (VP).

DEFINITION 1.2. Let $X_2 \subset X$. A function $h : X_2 \to R^1$ is said to be level-compact (level-bounded) on X_2 if the set $X_t = \{x \in X_2 : h(x) \leqslant t\}$ is compact (bounded) for any real number t.

The following function $\xi : Y \to R$ is very useful in the sequel:

$$\xi(y) = \min\{t \in R^1 : y \in te - C\}.$$

The proposition below summarizes some important properties of the function ξ.

PROPOSITION 1.1. [10]. *The function* ξ *is* (*stricktly*) *monotone* (*in the sense that* $y_1 \leqslant_C y_2$ *implies* $\xi(y_1) \leqslant \xi(y_2)$ *and* $y_2 - y_1 \in \mathrm{int}C$ *implies* $\xi(y_1) < \xi(y_2)$), *continuous, subadditive and positively homogenous on* Y. *Moreover,*

$$\xi(y) = \sup_{\lambda \in C^0} \lambda(y)/\lambda(e) \quad \forall y \in Y \tag{1}$$

and for any $y \in Y$ and $t \in R^1$, there holds

$$\xi(y + te) = \xi(y) + t.$$

Similarly, we can define $\eta : Z \to R^1$ by the formula:

$$\eta(z) = \min\{t \in R^1 : z \in tk^0 - K\}.$$

Clearly, the function η has the same properties as ξ stated in Proposition 1.1.

Consider the penalty problem:

$$(PVP_r^\alpha) \inf_{x \in X_1} F_\alpha(x, r) =: f(x) + r[d_{-K}(g(x))]^\alpha e,$$

where $d_{-K}(\cdot)$ is the distance function of a point to the set $-K$.

Denote by E_r' the set of infimum points of (PVP_r^α).

2. Convergence Analysis of Efficient Points

In this section, we show, under some conditions, that any efficient point of the cone constrained vector optimization problem can be approached by a sequence of efficient points of the penalty problems (PVP_r^α) as $r \to +\infty$ and any limit point of a sequence of approximate efficient solutions to the penalty problems is a weakly efficient solution of the original cone constrained vector optimization problem.

The following lemma is useful in the sequel.

LEMMA 2.1. *Let* $X_2 \subset X$ *be nonempty and closed and* $u : X_2 \to Y$ *be a vector-valued function such that* $\lambda(u)$ *is l.s.c on* X_2 *for any* $\lambda \in C^*$. *Assume that there exists* $x_0 \in X_2$ *such that the set* $A_0 = \{x \in X_2 : u(x) \leqslant_C u(x_0)\}$ *is compact. Then, u has an efficient solution on* X_2.

Proof. It is obvious that X_2 can be seen as a subspace of X. Let $y \in Y$. Then the set $\{x \in X_2 : u(x) \leqslant y\} = \cap_{\lambda \in C^*} \{x \in X_2 : \lambda(u(x)) \leqslant \lambda(y)\}$ is closed in X_2 by the l.s.c of $\lambda(u)$ for any $\lambda \in C^*$. Moreover, it can be shown by contradiction that if $\bar{x} \in A_0$ is an efficient solution of u on A_0, then it is also an efficient solution of u on X_2. The conclusion follows immediately from ([6], Corollary 3.1) (by replacing X and A with X_2 and A_0, respectively).

THEOREM 2.1. *Assume that*

$$Y^* = C^* - C^*. \tag{2}$$

and $\xi(f(x))$ is level-compact on

$$X_3 = \{x \in X_1 : g(x) \in \theta_0 k^0 - K\} \tag{3}$$

for some $\theta_0 > 0$. Further assume that there exist $r_0 > 0$ and $m_0 \in R^1$ such that

$$\xi(f(x)) + r_0 d^{\alpha}_{-K}(g(x)) \geqslant m_0, \ \forall x \in X_1 \tag{4}$$

Then, we have

(i) The set E of efficient points of (VP) is nonempty;
(ii) The set E of efficient points of (VP) coincide with the set E' of infimum points of (VP), i.e., $E = E'$;
(iii) Suppose that $0 < r_n \uparrow +\infty$. Then

$$E \subset w - \limsup_{n \to +\infty} E'_{r_n}$$

where

$$w - \limsup_{n \to +\infty} E'_{r_n} = \{y \in Y : \exists \text{ a subsequence } \{n_k\} \text{ of } \{n\}$$
$$\text{and } y_k \in E'_{r_{n_k}} \text{ such that } y_k \text{ weakly converges to } y\}.$$

Proof. (i) Suppose that $x_0 \in X_0$. Consider the set

$$\begin{aligned} X_4 &= \{x \in X_0 : f(x) \leqslant_C f(x_0)\} \\ &\subset \{x \in X_0 : \xi(f(x)) \leqslant \xi(f(x_0))\} \\ &\subset \{x \in X_3 : \xi(f(x)) \leqslant \xi(f(x_0))\}. \end{aligned}$$

By the level-compactness of $\xi(f)$ on X_3, we see that X_4 is nonempty and compact. Hence, by Lemma 2.1. $E \neq \emptyset$.

(ii) We need only to show that $E' \subseteq E$. Let $\bar{y} \in E'$. Then, there exists $\{x_n\} \subset X_0$ such that $f(x_n) \to \bar{y}$. Clearly, there exists $t_0 > 0$ such that $\xi(f(x_n)) \leqslant t_0$. By the level-compactness of $\xi(f)$ on X_3, we can assume without loss of generality $x_n \to \bar{x} \in X_0$. Thus, for any $\lambda \in C^*$,

$$\lambda(f(\bar{x})) \leqslant \lim_{n \to +\infty} \lambda(f(x_n)) = \lambda(\bar{y}).$$

It follows that $f(\bar{x}) \leqslant_C \bar{y}$. On the other hand, from $\bar{y} \in E'$, $\bar{y} - f(\bar{x}) \notin C \backslash \{0\}$. Hence, $\bar{y} = f(\bar{x})$. This proves $E' \subseteq E$.

(iii) Let $\bar{y} \in E$. Then, there exists $\bar{x} \in X_0$ such that $\bar{y} = f(x)$. Consider the set

$$X^n = \{x \in X_1 : F_\alpha(x, r_n) \leqslant_C F_\alpha(\bar{x}, r_n) = f(\bar{x})\}$$

It is obvious that $X^n \neq \emptyset$ since $\bar{x} \in X^n$. Now we show that there exists $n_0 > 0$ such that X^n is nonempty and compact whenever $n \geqslant n_0$. We need only to show that there exists $n_0 > 0$ such that $X^n \subset X_3$ when $n \geqslant n_0$ because $\xi(f(x)) \leqslant \xi(f(\bar{x})), \forall x \in X^n$ and $\xi(f)$ is level-compact on X_3. Note that $X^{n+1} \subset X^n, \forall n$. Suppose to the contrary that there exists a subsequence $\{n_k\}$ of $\{n\}$ such that $x_{n_k} \notin X_3, \forall k$. Then,

$$f(x_{n_k}) + r_{n_k} d^{\alpha}_{-K}(g(x_{n_k}))e \leqslant_C f(\bar{x}).$$

That is,

$$(f(x_{n_k}) + r_0 d^{\alpha}_{-K}(g(x_{n_k}))e) + (r_{n_k} - r_0) d^{\alpha}_{-K}(g(x_{n_k}))e \leqslant_C f(\bar{x}).$$

It follows that

$$(\xi(f(x_{n_k})) + r_0 d^{\alpha}_{-K}(g(x_{n_k}))) + (r_{n_k} - r_0) d^{\alpha}_{-K}(g(x_{n_k})) \leqslant \xi(f(\bar{x})).$$

By Equation (4), we have

$$m_0 + (r_{n_k} - r_0) d^{\alpha}_{-K}(g(x_{n_k})) \leqslant \xi(f(\bar{x})).$$

As a result,

$$d^{\alpha}_{-K}(g(x_{n_k})) \leqslant \frac{\xi(f(\bar{x})) - m_0}{r_{n_k} - r_0}.$$

Thus,

$$\lim_{k \to +\infty} d_{-K}(g(x_{n_k})) = 0,$$

implying $g(x_{n_k}) \in \theta_0 k^0 - K$ when k is large enough. That is, $x_{n_k} \in X_2$ when k is large. A contradiction arises. Hence, there exists $n_0 > 0$ such that X^n is nonempty and compact whenever $n \geqslant n_0$. Consequently, by Lemma 2.1, there exists $\bar{x}_n \in X^n$, which is an efficient solution to (PVP_{r_n}). Note that $\{\bar{x}_n\} \subset X^{n_0}, n \geqslant n_0$ and X^{n_0} is compact. We can assume without loss of generality that $\bar{x}_n \to x^*$. Then, from

$$f(\bar{x}_n) + r_n d^{\alpha}_{-K}(g(\bar{x}_n))e \leqslant_C f(\bar{x}), \tag{5}$$

We can show as arguing above that

$$d_{-K}(g(x^*)) = 0.$$

Thus, $x^* \in X_0$. Furthermore, from Equation (5), we have

$$f(\bar{x}_n) \leqslant f(\bar{x}).$$

Therefore,

$$\lambda(f(\bar{x}_n)) \leqslant \lambda(f(\bar{x})) \quad \forall \lambda \in C^*.$$

As a result,

$$\lambda(f(x^*)) \leqslant \liminf_{n \to +\infty} \lambda(f(\bar{x}_n)) \leqslant \lambda(f(\bar{x})) \quad \forall \lambda \in C^*.$$

Hence,

$$f(x^*) \leqslant f(\bar{x}).$$

On the other hand, $f(\bar{x}) \in E$. Consequently,

$$f(x^*) = f(\bar{x}). \tag{6}$$

For simplicity, let

$$\bar{y}_n = f(\bar{x}_n) + r_n d^{\alpha}_{-K}(g(\bar{x}_n))e. \tag{7}$$

Then, from Equation (5), we get

$$\lambda(f(\bar{x}_n)) \leqslant \lambda(\bar{y}_n) \leqslant \lambda(f(\bar{x})) \quad \forall \lambda \in C^*,$$

implying

$$\limsup_{n\to+\infty} \lambda(\bar{y}_n) \leqslant \lambda(f(\bar{x})), \quad \liminf_{n\to+\infty} \lambda(\bar{y}_n) \geqslant \lambda(f(x^*)).$$

This together with Equation (6) yields

$$\lim_{n\to+\infty} \lambda(\bar{y}_n) = \lambda(f(\bar{x}) = \lambda(f(x^*)) \quad \forall \lambda \in C^*.$$

As Equation (2) holds, so we further have

$$\lim_{n\to+\infty} l(\bar{y}_n) = l(f(\bar{x})) = l(f(x^*)) \quad \forall l \in Y^*.$$

That is, $\bar{y}_n$ weakly converges to $f(\bar{x})$. The proof is complete. □

Remark 2.1. (a) If C is a normal cone, then Equation (2) holds (see, e.g., [17]).

(b) If Y is of finite dimension and f is continuous, then we can show from Equations (5) and (7) that $\{\bar{y}_n\}$ is bounded. Otherwise, assume without loss of generality that $\bar{y}_n \to +\infty$ and $\bar{y}_n/\|\bar{y}_n\| \to y_0 \neq 0$. Then, from Equations (5) and (7), we deduce that

$$f(\bar{x}_n)/\|\bar{y}_n\| \leqslant_C \bar{y}_n/\|\bar{y}_n\| \leqslant_C f(\bar{x})/\|\bar{y}_n\|.$$

Passing to the limit as $n \to +\infty$, we obtain $0 \leqslant_C y_0 \leqslant_C 0$. By the pointedness of C, we conclude that $y_0 = 0$. A contradiction arises. So we can assume without loss of generality that $\bar{y}_n \to y^*$. Again, from Equations (5) and (7), we have

$$f(\bar{x}_n) \leqslant_C \bar{y}_n \leqslant_C f(\bar{x}).$$

Passing to the limit, we get

$$f(x^*) \leqslant_C y^* \leqslant_C f(\bar{x}).$$

This together with Equation (6) gives $y^* = f(\bar{x}) = \bar{y}$. Hence, statement (iii) of Theorem 2.1 holds without condition Equation (2).

(c) It can be easily checked that $\xi(f)$ is level-compact (level-bounded) on X_3 if and only if the set $\{x \in X_3 : f(x) \leqslant_C te\}$ compact (bounded) for any $t \in R^1$ because $f(x) \leqslant_C te$ if and only if $\xi(f(x)) \leqslant t$.

(d) Theorem 2.1 corresponds to Theorem 2.12 in [14].

If both X and Y are finite dimensional spaces and f is continuous, then Theorem 2.1 can be stated as the following result.

THEOREM 2.2. *Assume that X and Y are both finite dimensional, f and g are continuous on X_1. Further assume that Equation* (4) *holds for some $r_0 > 0$ and $m_0 \in R^1$ and $\xi(f)$ is level-bounded on X_3 (defined by Equation* (3)) *for some $\theta_0 > 0$.*

Then, for any sequence $0 < r_n \uparrow +\infty$,

$$\emptyset \neq E = E' \subset \limsup_{n\to+\infty} E_{r_n},$$

where E, E' are as in Theorem 2.1 and

$$\limsup_{n\to+\infty} E_{r_n} = \{y \in Y : \exists \text{ a subsequence } \{n_k\} \text{ of } \{n\} \text{ and } y_k \in E'_{r_{n_k}} \text{ such that } y_k \text{ converges to } y\}.$$

Proof. It follows immediately from Theorem 2.1 and statement (b) of Remark 2.1. □

Some sufficient conditions for the level-compactness of $\xi(f)$ on X_3 for some $\theta_0 > 0$ is stated in the following lemma.

LEMMA 2.2. *If one of the following conditions holds, then there exists $\theta_0 > 0$ such that $\xi(f)$ is level-compact on X_3 (defined by Equation* (3)).

(i) *$\xi(f)$ is level-compact on X_1;*

(ii) *$h(x) = \max\{\xi(f(x)), \eta(g(x)), 0\}$ is level-compact on X_1;*

(iii) *there exists $\theta_0 > 0$ such that X_3 is compact.*

Proof. The proof of the conclusion under condition (i) or (iii) is obvious. Now we prove the conclusion under condition (ii). Let $\theta > 0$ and $\alpha \in R^1$. Then, the set

$$A =: \{x \in X_1 : \xi(f(x)) \leqslant \alpha, g(x) \in \theta k^0 - K\} \subseteq \{x \in X_1 : \xi(f(x)) \leqslant \alpha, \eta(g(x)) \leqslant \theta\} \subseteq \{x \in X_1 : h(x) \leqslant \max\{\alpha, \theta\}\} =: B.$$

By (ii), the set B is compact. It is obvious that $\xi(f)$ is lower semicontinuous on X_1. It follows that A is closed. Hence, $A \subseteq B$ is compact. This shows that the θ_0 in X_3 can be any positive real number. □

Analogous to Lemma 2.2, we have the following result when X is a finite dimensional space.

LEMMA 2.3. *Assume that X is finite dimensional. If one of the following conditions holds, then there exists $\theta_0 > 0$ such that $\xi(f)$ is level-bounded on X_3 (defined by Equation (3)).*

(i) $\lim_{x \in X_1, \|x\| \to +\infty} \xi(f(x)) = +\infty$;

(ii) $\lim_{x \in X_1, \|x\| \to +\infty} \max\{\xi(f(x)), \eta(g(x)), 0\} = +\infty$;

(iii) *there exists $\theta_0 > 0$ such that X_3 is compact.*

Proof. It is known from [19] that a real-valued function f_0 defined on a subset X_1 of a finite dimensional space X is level-bounded if and only if $\lim_{x \in X_1, \|x\| \to +\infty} f_0(x) = +\infty$. The conclusion can be proved similarly to that of Lemma 2.1. □

The next two theorems follow directly from Theorem 2.1 and Lemma 2.2 as well as Theorem 2.2 and Lemma 2.3, respectively.

THEOREM 2.3. *Assume that Equation* (2) *and one of three conditions in Lemma* 2.2 *hold. Moreover, there exist $r_0 > 0$ and $m_0 \in R^1$ such that Equation* (4) *holds. Then the conclusion of Theorem* 2.1 *holds.*

THEOREM 2.4. *Assume that X and Y are both finite dimensional, f and g are continuous on X_1. Further assume that Equation* (4) *holds for some $r_0 > 0$ and $m_0 \in R^1$ and one of the three conditions in Lemma* 2.3 *holds. Then the conclusion of Theorem* 2.2 *holds.*

The next proposition shows that each limiting point of a sequence of approximate efficient solutions to the vector penalty problems is a weakly efficient solution of the original cone constrained vector optimization problem.

PROPOSITION 2.1. *Assume that f and g are both continuous on X_1 and $0 < \epsilon_n \to 0$ and $0 < r_n \to +\infty$. Suppose that each $\bar{x}_n \in X_1$ is an $\epsilon_n e$-efficient solution, i.e.,*

$$F_\alpha(x, r_n) - F_\alpha(\bar{x}_n, r_n) \notin -\epsilon_n e - C, \quad \forall x \in X_1. \tag{8}$$

Then, any limiting point of $\{\bar{x}_n\}$ is an efficient solution to (VP).

Proof. Suppose that $\bar{x}$ is a limiting point of $\{\bar{x}_n\}$ and assume without loss of generality that $\bar{x}_n \to \bar{x}$. Let $x_0 \in X_0$. Then, by Equation (8), we have

$$f(x_0) - f(\bar{x}_n) - r_n d^\alpha_{-K}(g(\bar{x}_n))e \notin -\epsilon_n e - C, \forall n. \tag{9}$$

Thus,

$$\xi(f(x_0) - f(\bar{x}_n)) - r_n d^\alpha_{-K}(g(\bar{x}_n)) \geqslant -\epsilon_n.$$

That is,

$$d^{\alpha}_{-K}(g(\bar{x}_n)) \leqslant \frac{\xi(f(x_0) - f(\bar{x}_n)) + \epsilon_n}{r_n}.$$

Passing to the limit as $n \to +\infty$, we obtain $d_{-K}(g(\bar{x})) = 0$. Hence, $g(\bar{x}) \in -K$ and $\bar{x} \in X_0$. Furthermore, from Equation (9), we deduce that

$$f(x_0) - f(\bar{x}_n) \notin -\epsilon_n e - C, \quad \forall n.$$

That is,

$$\xi(f(x_0) - f(\bar{x}_n)) > \epsilon_n, \quad \forall n.$$

Passing to the limit as $n \to +\infty$, we get

$$\xi(f(x_0) - f(\bar{x})) \geqslant 0.$$

It follows that $f(x_0) - f(\bar{x}) \notin -\mathrm{int}C$. By arbitrariness of $x_0 \in X_0$, we conclude that $\bar{x}$ is a weakly efficient solution to (VP).

3. Convergence Analysis via First-Order Optimality Conditions

Throughout this section, we assume that X is a Banach space, Y and Z are Hilbert spaces and Z is finite dimensional, C^* admits a weak*-compact base, $X_1 \subset X$ is nonempty, closed and convex, f and g are continuously differentiable on X.

In this section, we carry out convergence analysis via first-order necessary optimality conditions. Specifically, we show that any limit point of a sequence of points that satisfy first-order necessary optimality conditions of the penalty problems satisfies (KKT-type) first-order conditions of the orginal cone constrained vector optimization problem if the Mangasarian–Fromovitz constraint qualification holds at the limit point.

3.1. OPTIMALITY CONDITIONS FOR PENALTY PROBLEMS

In this subsection, we derive optimality conditions for a local weakly efficient solution to the penalty problem (PVP^{α}_r). Note that the penalty function is not differentiable when $\alpha \in (0, 1)$ or even not locally Lipschitz when $\alpha \in (0, 1)$. We will apply the Ekeland's variational principle for vector-valued functions developed in [4] to derive necessary optimality conditions for (PVP^{α}_r).

The following vector Ekeland's variational principle follows immediately from ([4], Corollary 2.1 and Lemma 2.3 (iii)).

LEMMA 3.1 *Let $h : X \to Y$ be a continuous vector-valued function and $X_2 \subset X$ be nonempty and closed. Let $\epsilon > 0$ be a scalar. Suppose that $\bar{x} \in X_2$ satisfies*

(a) $h(x) - h(\bar{x}) + \epsilon e \notin -C \setminus \{0\}$;

(b) *there exists a real number* m_0 *such that*

$$m_0 e \leqslant_C y - f(\bar{x}), \forall y \in (f(\bar{x}) - C) \cap f(X_2).$$

Then, there exists $x_\epsilon \in X_2$ such that
(i) $h(x_\epsilon) \leqslant_C h(\bar{x})$;
(ii) $\|x_\epsilon - \bar{x}\| \leqslant \sqrt{\epsilon}$;
(iii) $h(x) - h(x_\epsilon) + \sqrt{\epsilon}\|x - x_\epsilon\| e \notin -C, \forall x \in X_2 \setminus \{x_\epsilon\}$.

Let $x \in X_1$. Denote by $T_{X_1}(x)$ and $N_{X_1}(x)$ the tangent cone and normal cone of X_1 at the point x, respectively.

The next lemma is useful in deriving optimality conditions for (PVP_r^α).

LEMMA 3.2 [11]. *Let* $h : X \to Y$ *be a locally Lipschitz vector-valued function. Suppose that* $\bar{x}$ *is a local weakly efficient solution to the vector optimization problem:* $\min_{x \in X_1} h(x)$. *Then, there exists* $\lambda \in C^0$ *such that* $0 \in \partial(\lambda(h))(\bar{x}) + N_{X_1}(\bar{x})$, *where* ∂ *is the Clarke subdifferential.*

PROPOSITION 3.1. *Let* $\bar{x}_r \in X_1$ *be a local weakly efficient solution to* (PVP_r^α). *Suppose that the following condition holds:*

$$(C) \quad \exists \bar{d}_r \in T_{X_1}(\bar{x}_r) \text{such that } g(\bar{x}_r) + \nabla g(\bar{x}_r)\bar{d}_r, \in -\text{int}C. \tag{10}$$

Then, there exist $\lambda_r \in C^* \setminus \{0\}, \mu_r \in Z^*$ with

$$\mu_r = \alpha r \lambda_r(e)[d_{-K}^{\alpha-2}(g(\bar{x}_r))](g(\bar{x}_r) - P_{-K}(g(\bar{x}_r)))\nabla g(\bar{x}_r), \text{ if } \bar{x}_r \notin -X_0, \tag{11}$$

$$\mu_r \in K^*, \mu_r(g(\bar{x}_r)) = 0, \text{ if } \bar{x}_r \in X_0, \tag{12}$$

where $P_{-K}(y)$ is the projection of y onto $-K$, such that

$$0 \in \lambda_r(\nabla f(\bar{x}_r)) + \mu_r(\nabla(g(\bar{x}))) + N_{X_1}(\bar{x}_r). \tag{13}$$

Proof. Since $\bar{x}_r$ is a local weakly efficient solution to (PVP_r^α), there exists $\delta > 0$ such that $\bar{x}_r$ is a weakly efficient solution of the vector-valued function $F_\alpha(x, r)$ on

$$U_\delta = \{x \in X_1 : \|x - \bar{x}\| \leqslant \delta\}.$$

That is,

$$F_\alpha(x, r) - F_\alpha(\bar{x}_r, r) \notin -\text{int}C, \forall x \in U_\delta. \tag{14}$$

As $F_\alpha(x, r)$ is continuous (in x) on X_1, we can choose $\delta > 0$ such that, for some real number γ,

$$\gamma e \leqslant_C F_\alpha(x, r) - F_\alpha(\bar{x}_r, r), \forall x \in U_\delta. \tag{15}$$

Let

$$s_n(x) = f(x) + r[d_{-K}^2(g(x)) + 1/n^2]^{\alpha/2} e.$$

Then, It is easy to see that there exists $0 < \epsilon_n' \downarrow 0$ such that

$$0 \leqslant_C s_n(x) - F_\alpha(x, r) \leqslant_C \epsilon_n e, \forall x \in X_1. \tag{16}$$

It follows from Equations (14) and (16) that

$$s_n(x) - s_n(\bar{x}) + \epsilon_n' e \notin -\operatorname{int} C, \forall x \in U_\delta.$$

Let $\epsilon_n = 2\epsilon_n'$. Then, showing by contradiction, we have

$$s_n(x) - s_n(\bar{x}_r) + \epsilon_n e \ \notin -C, \forall x \in U_\delta. \tag{17}$$

Clearly, $0 < \epsilon_n \downarrow 0$ as $n \to +\infty$. Assume without loss of generality that

$$\sqrt{\epsilon_n} < \delta, \forall n. \tag{18}$$

Taking $X_2 = U_\delta$, the combination of Equations (15), (16) and (17) implies that all the conditions of Lemma 3.1 hold (with the function h replaced by s_n). By Lemma 3.1, there exists $x_n \in U_\delta$ such that

$$\|x_n - \bar{x}_r\| \leqslant \sqrt{\epsilon_n}, \tag{19}$$

$$s_n(x) - s_n(x_n) + \sqrt{\epsilon_n}\|x_n - \bar{x}_r\| e \notin -C, \forall x \in U_\delta \setminus \{x_n\}. \tag{20}$$

By Lemma 3.2 and the chain rule for the Clarke subdifferential, there exists $\lambda_n' \in C^0$ such that

$$0 \in \partial(\lambda_n'(s_n))(x_n) + \sqrt{\epsilon_n}\lambda_n'(e)B + N_{U_\delta}(x_n), \tag{21}$$

where $B = \{v \in X^* : \|v\| \leqslant 1\}$. By Equations (18) and (19), we have

$$N_{U_\delta}(\bar{x}_n) = N_{X_1}(x_n). \tag{22}$$

Note that

$$\lambda_n'(s_n(x)) = \lambda_n'(f(x)) + r\lambda_n'(e)[d_{-K}^2(g(x)) + 1/n^2]^{\alpha/2},$$

$d_{-K}^2(\cdot)$ is $C^{1,1}$ (see, e.g., [1,9]) and

$$\nabla d_{-K}^2(y) = 2(y - P_{-K}(y)), \forall y \in Y.$$

Hence, s_n is continuously differentiable, and

$$\begin{aligned}
&\nabla \lambda_n'(s_n)(x_n) \\
&= \lambda_n'(\nabla f(x_n)) + r\alpha/2\lambda_n'(e)[d_{-K}^2(g(x_n)) + 1/n^2]^{\alpha/2-1}\nabla d_{-K}^2(g)(x_n) \\
&= \lambda_n'(\nabla f(x_n)) + r\alpha/\lambda_n'(e)[d_{-K}^2(g(x_n)) + 1/n^2]^{\alpha/2-1}(g(x_n) \\
&\quad - P_{-K}(g(x_n)))(\nabla g(x_n)).
\end{aligned} \tag{23}$$

Recall that C^* admits a weak*-compact base and that $\lambda_n' \in C^0$. Suppose that $C_1 \subset C^*$ is a weak*-compact base of C^*. Then, there exist $0 < t_n \in R_+^1$ and $\lambda_n'' \in C_1$ such that $\lambda_n' = t_n\lambda_n''$. Assume without loss of generality that $w^* - \lim_{n\to+\infty}\lambda_n'' = \lambda_0 \in C_1$, where w^*-lim denotes weak * convergence. Obviously, $\lambda_0 \neq 0$. Furthermore, $\lim_{n\to+\infty} t_n = t_0 > 0$. Thus, we can assume without loss of generality that

$$w^* - \lim_{n\to+\infty} \lambda_n' = t_0\lambda_0 =: \lambda_r \in C^*\backslash\{0\}. \tag{24}$$

Let

$$\mu_n' = r\alpha\lambda_n'(e)[d_{-K}^2(g(x_n)) + 1/n^2]^{\alpha/2-1} \quad (g(x_n) - P_{-K}(g(x_n))). \tag{25}$$

Then, from Equations (21)–(23), we deduce that for each n, there exists $\xi_n \in B$ such that

$$\lambda_n'(\nabla f(x_n)) + \mu_n'(\nabla g(x_n)) + \sqrt{\epsilon_n}\lambda_n'(e)\xi_n \in -N_{X1}(x_n). \tag{26}$$

Consider the following two cases.

(a) $\bar{x}_r \notin X_0$. Let

$$\mu_r = \lim_{n\to+\infty} \mu_n',$$

giving rise to Equation (11). Taking the limit in Equation (26) as $n \to +\infty$, we obtain Equation (13).

(b) $\bar{x} \in X_0$. We first show that $\{\|\mu_n'\|\}$ is bounded. Otherwise, assume without loss of generality that $\|\mu_n'\| \to +\infty$ and

$$\mu_n'/\|\mu_n'\| \to \mu_r' \neq 0 \tag{27}$$

(which is true since Z is finite dimensional). Dividing Equation (26) by $\|\mu_n'\|$ and passing to the limit as $n \to +\infty$, we obtain

$$\mu_r'(\nabla g(\bar{x}_r)) \in -N_{X1}(\overline{x}_r). \tag{28}$$

Note from Equation (25) that

$$\frac{\mu_n'}{\|\mu_n'\|} = \frac{g(x_n) - P_{-K}(g(x_n))}{\|g(x_n) - P_{-K}(g(x_n))\|}. \tag{29}$$

Thus, for any $v \in -K$

$$(g(x_n) - P_{-K}(g(x_n)))(v) = (g(x_n) - P_{-K}(g(x_n)))(v - P_{-K}(g(x_n))) \leqslant 0$$

by the basic properties of a projection map onto a closed convex cone in a Hilbert space. Thus, $(\mu_n'/\|\mu_n'\|)(v) \leqslant 0$. As a result, $\mu_r'(v) \leqslant 0$. Hence

$$\mu_r' \in K^*. \tag{30}$$

Note also that

$$\begin{aligned}&(g(x_n) - P_{-K}(g(x_n)))(g(x_n))\\&= (g(x_n) - P_{-K}(g(x_n)))(g(x_n) - P_{-K}(g(x_n))) \geqslant 0.\end{aligned}$$

This combined with Equations (29) and (27) yields $\mu_r'(g(\bar{x}_r)) \geqslant 0$. On the other hand, from $x \in X_0$ and Equation (30), we have $\mu_r'(g(\bar{x}_r) \leqslant 0$. Hence,

$$\mu_r'(g(\bar{x}_r)) = 0. \tag{31}$$

Moreover, the combination of Equations (10) and (31) yields

$$\mu_r'(\nabla g(\bar{x}_r))(\overline{d}_r)) = \mu_r'(g(\bar{x}_r) + \nabla g(\bar{x}_r)(\overline{d}_r)) < 0. \tag{32}$$

On the other hand, Equation (28) and the fact that $\overline{d}_r \in T_{X1}(\bar{x}_r)$imply that

$$\mu_r'(\nabla g(\bar{x}_r))(\overline{d}_r)) \geqslant 0,$$

contradicting Equation (32). Hence, $\{\mu_n'\}$ is bounded. Assume without loss of generality that $\mu_n' \to \mu_r$. Taking the limit in Equation (26) as $n \to +\infty$, we get Equation (13). Arguing as above for $\mu_r' \in K^*$ and $\mu_r'(g(\overline{x}_r)) = 0$, we can show that Equation (12) holds. The proof is complete. □

3.2. CONVERGENCE OF STATIONARY POINTS OF PENALTY PROBLEMS

In this subsection, we show under some conditions that any limit point of the points that satisfy the conditions stated in Proposition 3.1 is a KKT point of the original (VP).

DEFINITION 3.1. Let $\bar{x} \in X_0$. $\bar{x}$ is said to satisfy the KKT condition if there exists $(\lambda, \mu) \in C^0 \times K^*$ such that

$$0 \in \lambda(\nabla f(\bar{x})) + \mu(\nabla g(\overline{x})) + N_{X1}(\bar{x}), \tag{33}$$

$$\mu(g(\bar{x})) = 0, \tag{34}$$

DEFINITION 3.2. Let $\bar{x} \in X_0$. We say that the Mangasarian–Fromovitz constraint qualification (MFCQ) holds at $\bar{x}$ if there exists $d \in T_{X1}(\bar{x})$ such that $g(\bar{x}) + \nabla g(\bar{x})(d) \in -\mathrm{int}K$.

Now we present the main results of this section.

THEOREM 3.1. *Let $0 < r_n \uparrow +\infty$ and $\bar{x}_n \in X_1$. Suppose that there exists $M \in R^1$ such that*

$$f(\bar{x}_n) + r_n d^{\alpha}_{-K}(g(\bar{x}_n))e \leqslant_C Me. \tag{35}$$

Suppose that $\bar{x}$ is limit point of $\{\bar{x}_n\}$. Then, $\bar{x} \in X_0$. Assume that MFCQ holds at $\bar{x}$. Then condition (C) (with r replaced by r_n in Proposition 3.1) holds when n sufficiently large. Suppose that the first-order necessary condition Equation (13) for (PVP^{α}_{rn}) holds when n is large. Then $\bar{x}$ is a KKT point of (VP).

Proof. Assume without loss of generality that $\bar{x}_n \to \bar{x}$. The assertion that $\bar{x} \in X_0$ follows immediately from Equation (35). It is obvious that condition (C) holds when n is sufficiently large if MGCQ holds at $\bar{x}$. By the first order necessary condition for (PVP^{α}_{rn}), there exist $\lambda_n' \in C^*\backslash\{0\}$ and μ_n' with

$$\mu'_n = \alpha r_n \lambda'_n(e)[d_{-K}^{\alpha-2}(g(\bar{x}_n))](g(\bar{x}_n) - P_{-K}(g(\bar{x}_n)))\nabla g(\bar{x}_n), \text{ if } \bar{x}_n \notin -X_0,$$
$$\mu'_n \in K^*, \mu'_n(g(\bar{x}_n)) = 0 \text{ if } g(\bar{x}_n) \in X_0$$

such that

$$0 \in \lambda'_n(\nabla f(\bar{x}_n)) + \mu'_n(\nabla(g(\bar{x}_n))) + N_{X_1}(\bar{x}_n).$$

Let $\lambda_n = \lambda'_n/\|\lambda'_n\|$ and $\mu_n = \mu_n/\|\lambda'_n\|/$. Then, $\lambda_n \in C^0$ and

$$\mu_n = \alpha r_n \lambda_n(e)[d_{-K}^{\alpha-2}(g(\bar{x}_n))](g(\bar{x}_n) - P_{-K}(g(\bar{x}_n)))\nabla g(\bar{x}_n), \text{if } \bar{x}_n \notin -X_0, \tag{36}$$

$$\mu_n \in K^*, \quad \mu_n(g(\bar{x}_n)) = 0, \text{ if } g(\bar{x}_n) \in X_0 \tag{37}$$

and

$$0 \in \lambda_n(\nabla f(\bar{x}_n)) + \mu_n(\nabla(g(\bar{x}_n))) + N_{X_1}(\bar{x}_n). \tag{38}$$

Arguing as in the proof of Proposition 3.1, we can show that $\{\mu_n\}$ is bounded and that there exists a subsequence of $\{\lambda_n\}$ which converges to $\lambda' \in C^0\backslash\{0\}$. Assume without loss of generality that $\lambda_n \to \lambda'$ and $\mu_n \to \mu'$. Taking the limit in Equation (38) as $n \to +\infty$, we obtain

$$0 \in \lambda'(\nabla f(\bar{x})) + \mu'(\nabla(g(\bar{x}))) + N_{X_1}(\bar{x}). \tag{39}$$

Consider the following two cases.

(i) There exists a subsequence $\{n_k\}$ of $\{n\}$ such that $x_{nk} \in X_0, \forall k$. Then, from Equation (37) (with n replaced by n_k), we have $\mu' \in K^*$ and $\mu'(g(\bar{x})) = 0$.

(ii) $\bar{x}_n \notin X_0$ when $n \geqslant n_0$ for some $n_0 > 0$. Then, μ_n is defined by Equation (36). Note that $(g(\bar{x}_n) - P_{-K}(g(\bar{x}_n)))(g(\bar{x}_n)) = 0$. It follows that $\mu_n(g(\bar{x}_n)) = 0$. Thus, $\mu'(g(\bar{x})) = 0$. Note also that for any $v \in -K$, $(g(\bar{x}_n) - P_{-K}(g(\bar{x}_n)))(v) = (g(\bar{x}_n) - P_{-K}(g(\bar{x}_n)))(v - g(\bar{x}_n)) \leqslant 0$. It follows that $\mu_n \in K^*$. Hence, $\mu' \in K^*$.

Let $\lambda = \lambda'/\|\lambda'\|$ and $\mu = \mu'/\|\lambda'\|$. Then ,$\lambda \in C^0, \mu \in K^*$ and Equation (34) holds. Moreover, from Equation (39), we get Equation (33). Hence, $\bar{x}$ is a KKT Point of (VP).

References

1. Asplund, E. (1979), Differentiability of the metric projection in finite dimensional Euclidean space, *Proceedings of the American Mathematical Society* 38, 229–238.
2. Burke, J.V. (1991), Calmness and exact penalization, *SIAM Journal on Control and Optimization* 494–497.
3. Burke, J.V. (1991), An exact penalization viewpoint of constrained optimization, *SIAM Journal on Control and Optimization* 29, 968–998.

4. Chen, G.Y. and Huang, X.X. (1998), A unified approach to the existing three types of variational principles for vector-valued functions, *Mathematical Methods of Operations Research* 48, 349–357.
5. Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis,* John-Wiley & Sons, New-York.
6. Corley, H.W. (1980), An existence result for maximizations with respect to cones, *Journal of Optimization Theory and Applications* 31, 277–281.
7. Fiacco, A. and McCormic, G. (1968), *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* John Wiley & Sons, New York.
8. Fletcher, R. (1987), *Practical Methods of Optimization*, John Wiley & Sons, New York.
9. Holmes, R.B. (1972), *A Course on Optimization and Best Approximation,* Springer, New York
10. Huang, X.X. (2000), Extended well-posedness properties of vector optimization problems, *Journal of Optimization Theory and Application* 106, 165–182.
11. Huang, X.X. (2003), Optimality conditions and approximate optimality conditions for locally Lipschitz vector optimization, *optimization* 51, 309–321.
12. Huang, X.X., Teo, K.L. and Yang, X.Q. (2003), Calmness and exact penalization for vector optimization problems with cone constraints, Preprint, Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong.
13. Huang, X.Q. and Yang, X.Q. (2001), Asymptotic analysis of a class of nonlinear penalization methods for constrained multiobjective optimization, *Nonlinear Analysis, TMA,* 5573–5584.
14. Huang, X.X, and Yang, X.Q. (2002), Nonlinear Lagrangian for multiobjective optimization problems and applications to duality and exact penalization, *SIAM Journal of Optimization* 13, 675–692.
15. Haung, X.X. and Yang, X.Q. (2003), Convergence analysis of a class of nonlinear penalization method via first order necessary optimality conditions, *Journal of Optimization Theory and Applications* 116, 311–332.
16. Huang, X.X. and Yang, X.Q. (2003), A Unifield augmented Lagrangian approach to duality and exact penalization, *Mathematics of Operation Research* 28, 533–552.
17. Isac, G. (1996), The Ekeland's principle and the Pareto ϵ-efficiency. In: Tamiz, M.(ed.), *Multiobjective programming and goal programming, theory and applications,* Lecture Notes in Economics and Mathematical Systems, Vol. 432, Springer-Verlag.
18. Luo, Z.Q. Pang, J.S. and Ralph, D. (1997), *Mathematical Programs with Equilibrium Constraints,* Cambridge University Press.
19. Rockafeller, R.T. and Wets, R.J.-B. (1988), *Variational Analysis,* Springer New york.
20. Ruan,G.Z. and Huang, X.X. (1992),Weak calmness and weak stability of multiobjective programming and its penalty functions, *Systems Sciences and Mathematical Sciences* 148–157, (in Chinese).
21. Rubinov, A.M., Glover, B.M. and Yang, X.Q. (1999), Decreasing functions with applications to penalization *SIAM Journal of Optimization* 289–313.
22. Rubinov, A. and Yang, X.Q. (2003), *Lagrang-Type Functions in Constrained Non-convex Optimization*, Kluwer Academic Publishers.
23. White, D.J. (1984), Multiobjective programming and penalty functions, *Journal of Optimization Theory and Applications* 43, 583–599.
24. Yang, X.Q. and Huang, X.X. (2001), A nonlinear Lagrangian approach to constrained optimization problems, *SIAM Journal of Optimization* 11, 1119–1144.