Discussiones Mathematicae
Differential Inclusions, Control and Optimization 30 (2010) 101–120

PROJECTION METHOD WITH LEVEL CONTROL IN CONVEX MINIMIZATION

Robert Dylewski

Faculty of Mathematics, Computer Science
and Econometrics, University of Zielona Góra
65–516 Zielona Góra, ul. Prof. Z. Szafrana 4a, Poland

e-mail: r.dylewski@wmie.uz.zgora.pl

Abstract

We study a projection method with level control for nonsmoooth convex minimization problems. We introduce a changeable level parameter to level control. The level estimates the minimal value of the objective function and is updated in each iteration. We analyse the convergence and estimate the efficiency of this method.

Keywords: projection method, convex nondifferentiable minimization, level control.

2000 Mathematics Subject Classification: 65K05, 90C25.

1. Introduction

We consider the convex minimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in D, \end{array} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex (not necessarily differentiable) function and $D \subset \mathbb{R}^n$ is a nonempty, convex and compact subset. Then the solution set

$$M = \operatorname*{Argmin}_{x \in D} f(x) = \{z \in D : f(z) \leq f(x) \text{ for all } x \in D\}$$

is nonempty, i.e., f attains its minimum $f^* = \min\{f(x) : x \in D\}$.

We suppose that for any $x \in D$ we can evaluate the objective value $f(x)$ and a single subgradient $g_f(x)$, and that for any $x \in \mathbb{R}^n$ we can evaluate the metric projection $P_D(x)$ of x onto D.

We use the following notation:

$x = (\xi_1, \dots, \xi_n)^\top$ – an element of $\mathbb{R}^n$,

x_k – kth element of a sequence (x_k),

$\langle x, y \rangle = \sum_{i=1}^n \xi_i \eta_i$ – the standard scalar product of vectors $x, y \in \mathbb{R}^n$,

$\|x\| = \sqrt{\langle x, x \rangle}$ – the Euclidean norm of a vector $x \in \mathbb{R}^n$,

$S(h, \alpha) = \{x \in \mathbb{R}^n : h(x) \leq \alpha\}$ – the sublevel set of a function h with a level α,

$S'(h, \alpha) = \{x \in \mathbb{R}^n : h(x) < \alpha\}$,

$\partial f(x) = \{g \in \mathbb{R}^n : f(y) - f(x) \geq \langle g, y - x \rangle, y \in \mathbb{R}^n\}$ – the subdifferential of a function f at x,

$g_k = g_f(x_k)$ – a subgradient of f at $x_k \in \mathbb{R}^n$ (any element of $\partial f(x_k)$),

$f_k(\cdot) = \langle g_k, \cdot - x_k \rangle + f(x_k)$ – a linearization of f at x_k,

$\check{f}_k = \max_{1 \leq i \leq k} f_i$ – the best model (lower bound) of f,

$\check{f}_k^* = \min_{x \in D} \check{f}_k(x)$,

$d(x, C) = \inf_{z \in C} \|z - x\|$ – the distance of x to the subset C,

$\operatorname{diam}(C) = \sup_{x,y \in C} \|y - x\|$ – the diameter of subset C,

$P_C(x) = \operatorname{argmin}_{y \in C} \|y - x\|$ – the metric projection of x onto a closed, convex subset $C \subset \mathbb{R}^n$.

We study the projection method, with level control for problem (1), of the form

$$x_1 \in D \text{ – arbitrary} \qquad (2)$$
$$x_{k+1} = P_D(x_k + \lambda_k t_k),$$

where:

- $\lambda_k \in (0, 2)$ is so called relaxation parameter,
- vector t_k has the form

$$t_k = P_{\bigcap_{i \in L_k} S(f_i, \alpha_k)} x_k - x_k, \qquad (3)$$

- $L_k \subset \{1, 2, \dots, k\}$ is a subset of saved linearization,

- $\alpha_k = (1-\nu_k)\overline{\alpha}_k + \nu_k \underline{\alpha}_k$ denotes the current level (an approximation of the minimal value f^* of the objective function f),
- $\nu_k \in (0,1]$ is a level parameter,
- $\overline{\alpha}_k = \min_{1\le i\le k} f(x_i)$ is an upper bound of f^*,
- $\underline{\alpha}_k \le f^*$ is a lower bound of f^* which is updated in each iteration.

Additionally, we assume that we know:

– an initial lower bound $\underline{\alpha}_1$ of f^*,

– an upper bound R of the distance of the starting point x_1 to the solution set M, $R \ge d(x_1, M)$.

Remark 1. The presented method is a genaralization of the following methods.

a) Let f^* be known. If we set $\nu_k = 1$ and $\underline{\alpha}_k = f^*$, then $\alpha_k = f^*$. If additionaly, $L_k = \{k\}$, then $\bigcap_{i\in L_k} S(f_i,\alpha_k) = S(f_k, f^*)$ and $t_k = -\frac{f(x_k)-f^*}{\|g_k\|}\frac{g_k}{\|g_k\|}$. We obtain the Polyak subgradient projection method [8].

b) Let $\nu_k = \nu \in (0,1)$. If $L_k = \{k\}$, then $\bigcap_{i\in L_k} S(f_i,\alpha_k) = S(f_k,\alpha_k)$ and $t_k = -\frac{f(x_k)-\alpha_k}{\|g_k\|}\frac{g_k}{\|g_k\|}$. We obtain the variable target value subgradient method of Kim-Ahn-Cho [5].

c) Let $\underline{\alpha}_k = \check{f}_k^*$ (of course $\underline{\alpha}_k \le f^*$) and $\nu_k = \nu \in (0,1)$. If $L_k = \{1,2,\ldots,k\}$, then $\bigcap_{i\in L_k} S(f_i,\alpha_k) = S(\check{f}_k,\alpha_k)$. We obtain the level method of Lemaréchal-Nemirovskii-Nesterov [7].

d) Let $\nu_k = \nu \in (0,1)$ and let $L_k \subset \{1,2,...,k\}$ such that $k \in L_k$. Then $\bigcap_{i\in L_k} S(f_i,\alpha_k) = S(f_{L_k},\alpha_k)$, where $f_{L_k} = \max_{i\in L_k} f_i$. We obtain the subgradient projection method with level control proposed by Kiwiel [6].

e) Let $\nu_k = \nu \in (0,1)$. If $L_k \subset \{1,2,\ldots,k\}$ is such that the system of subgradients $\{g_i : i \in L_k\}$ is linearly independent and generates an obtuse cone, then $\bigcap_{i\in L_k} S(f_i,\alpha_k) = S(f_{L_k},\alpha_k)$ for model $f_{L_k} = \max_{i\in L_k} f_i$. We obtain the method of projection with level control and obtuse cone selection proposed by Cegielski [2].

f) Let $\nu_k = \nu \in (0,1)$. Let $L_k \subset \{1,2,\ldots,k\}$ be such that the system of subgradients $\{g_i : i \in L_k\}$ is obtained from so called residual selection model. We have the method of projection with level control and residual selection studied in [3] and [4].

In Section 2 we present a general iterative scheme for the considered projection method with level control. In Section 3 we analyse the convergence of the method. In the last section we estimate the efficiency of the method.

2. Projection method with level control

Now we formulate the general projection method with level control.

Recall that the point $x_\varepsilon \in D$ is an ε-optimal solution of problem (1) if it satisfies the following condition:

$$f(x_\varepsilon) \leq f(x) + \varepsilon \text{ for all } x \in D. \tag{4}$$

Let (x_k) be a sequence generated by the following iterative scheme, which is a modification of the schemes presented in [2, Iterative Scheme 2], [6, Algorithm 2.2].

Iterative Scheme 2. *(Projection method with level control)*

Step 0. (*Initialization*)

0.1 Choose: $x_1 \in D$ (*starting point*), $\varepsilon \geq 0$ (*optimality tolerance*), $\underline{\lambda}, \overline{\lambda} \in (0,2)$ such that $\underline{\lambda} \leq \overline{\lambda}$ (*lower and upper bounds of the relaxation parameter*), $\underline{\nu}, \overline{\nu} \in (0,1)$ such that $\underline{\nu} \leq \overline{\nu}$ (*lower and upper bounds of the level parameter*), $R \geq d(x_1, M)$ (*upper bound of the distance of the starting point x_1 to the solution set*), $\underline{\alpha}_1 \in (-\infty, f^*]$ (*initial lower bound of f^**), $\overline{\alpha}_0 \in (f(x_1), +\infty)$ (*initial upper bound of f^**), $m \geq 1$ (*number of saved linearizations*).

0.2 Set: $k = 1$ (*iterations counter*), $l = 0$ (*counter of updates of the lower bound $\underline{\alpha}_k$*), $r_1 = 0$ (*initial distance parameter*), $\overline{x}_1 = x_1$.

Step 1. (*Objective evaluations*)

Calculate $f(x_k)$ and $g_k \in \partial f(x_k)$.

Step 2. (*Upper bound update*)

If $f(x_k) < \overline{\alpha}_{k-1}$ set $\overline{\alpha}_k = f(x_k)$ and $\overline{x}_k = x_k$.

Otherwise, set $\overline{\alpha}_k = \overline{\alpha}_{k-1}$ and $\overline{x}_k = \overline{x}_{k-1}$.

Step 3. (*Stopping criterion*)

3.1 If $\overline{\alpha}_k - \underline{\alpha}_k \leq \varepsilon$, then terminate ($\overline{x}_k$ *is an* ε*-optimal solution*).

3.2 If $\|g_k\|\, R \leq \varepsilon$, then terminate ($x_k$ *is an* ε*-optimal solution*).

Step 4. (*Level update*)

4.1 Choose $\nu_k \in [\underline{\nu}, \overline{\nu}]$.

4.2 Set $\alpha_k = (1-\nu_k)\overline{\alpha}_k + \nu_k \underline{\alpha}_k$.

Step 5. (*Update of saved linearizations of* f)

Set $J_k = \{k-m+1, \ldots, k\}$.

Step 6. (*Selection of linearizations*)

6.1 Choose an appropriate subset $L_k \subset J_k$ such that $k \in L_k$.

6.2 If the equality $S_k := \bigcap_{i\in L_k} S(f_i, \alpha_k) = \emptyset$ is detected, then go to Step 10 (*level* α_k *is too low*).

Step 7. (*Projection*)

7.1 Construct $t_k = P_{S_k}(x_k) - x_k$.

7.2 Choose $\lambda_k \in [\underline{\lambda}, \overline{\lambda}]$.

7.3 Evaluate $z_k = x_k + \lambda_k t_k$.

7.4 Evaluate $z'_k = P_D z_k$ and $q_k = z'_k - z_k$.

Step 8. (*Inconsistency detection*)

8.1 Set:
$r'_k = r_k + \lambda_k(2-\lambda_k)\|t_k\|^2 + \|q_k\|^2$,
$r''_k = r_k + \|t_k\|^2$.

8.2 If
$r'_k > R^2 - (R - \|z'_k - x_{k'+1}\|)^2$ or
$r''_k > R^2 - (R - \|x_k + t_k - x_{k'+1}\|)^2$,
where k' is the last iteration in which Step 10 was executed (initial $k' = 0$), then go to Step 10 (*level* α_k *is too low*).

Step 9. (*Approximation update*)

9.1 Set $x_{k+1} = z'_k$.

9.2 Set $r_{k+1} = r'_k$.

9.3 Increase k by 1 and go to Step 1.

Step 10. (*Lower bound update*)

10.1 Set $\underline{\alpha}_{k+1} = \alpha_k$, $\overline{\alpha}_{k+1} = \overline{\alpha}_k$ and $\overline{x}_{k+1} = \overline{x}_k$.

10.2 Set $r_{k+1} = 0$.

10.3 Set $x_{k+1} = \overline{x}_k$.

10.4 Increase k and l by 1 and go to Step 3.

Steps 6 and 7 were discussed in detail in [2, 3, 6] and [4].

Remark 3.

a) By the definition of subgradient we have inequality

$$f(x) \geq f(x_1) - \|g_1\|\, R' \tag{5}$$

for all $x \in D$, where $R' \geq \operatorname{diam}(D)$. Indeed, the subgradient g_1 of f at x_1 satisfies the inequality

$$f(x) \geq f(x_1) - \langle g_1, x_1 - x \rangle .$$

By the Schwarz inequality and inequality $R' \geq \operatorname{diam}(D)$, we obtain (5) for all $x \in D$. If we do not know a better initial lower bound $\underline{\alpha}_1$ of f^*, then we can take

$$\underline{\alpha}_1 = f(x_1) - \|g_1\|\, R'.$$

b) From the equalities in Steps 2 and 10.1 we have

$$\overline{x}_k = \operatorname*{argmin}_{1 \leq i \leq k} f(x_i).$$

c) If $L_k = \{k\}$ in Step 6.1, then Iterative Scheme 2 assigns the vector t_k such as in the method of Kim-Ahn-Cho [5]. In this case we have

$$S_k = \{x \in \mathbb{R}^n : f_k(x) = \langle g_k, x - x_k \rangle + f(x_k) \leq \alpha_k\}$$

and

$$t_k = -\frac{(f(x_k) - \alpha_k) g_k}{\|g_k\|^2}.$$

Furthermore, if $k \in L_k$ and $S_k = \bigcap_{i \in L_k} S(f_i, \alpha_k) \neq \emptyset$, then we obtain

$$\|t_k\| \geq \frac{f(x_k) - \alpha_k}{\|g_k\|}. \tag{6}$$

We denote

$$f_L = \max_{i \in L} f_i, \tag{7}$$

where $L \subset \{1, 2, \ldots, k\}$. Since $f_i \leq f, i \in L$ and $f_i(x_i) = f(x_i)$ for $i \in L$, we have $f(x) \geq f_L(x)$ for all $x \in \mathbb{R}^n$ and $f(x_i) = f_L(x_i)$ for $i \in L$.

The following lemmas explain why we have to go to Step 10 when the condition in Step 6.2 is satisfied.

Lemma 4. *If functions $h, f : \mathbb{R}^n \to \mathbb{R}$ are such that $h \leq f$, then $S(f, \alpha) \subset S(h, \beta)$ for $\alpha, \beta \in \mathbb{R}$ such that $\alpha \leq \beta$.*

Proof. Let $y \in S(f, \alpha)$. Therefore, $f(y) \leq \alpha$. By the assumption of the lemma, we have $h(y) \leq f(y)$. Consequently $h(y) \leq \alpha \leq \beta$ and $y \in S(h, \beta)$. ∎

Lemma 5. *If $\beta < \alpha$ then $S(f, \beta) \subset S'(f, \alpha) \subset S(f, \alpha)$.*

Proof. Let $y \in S(f, \beta)$, then $f(y) \leq \beta$. By the assumption of the lemma, we obtain $f(y) \leq \beta < \alpha$. Hence, $y \in S'(f, \alpha)$ and consequently $y \in S(f, \alpha)$. ∎

Lemma 6. *Let function $h : \mathbb{R}^n \to \mathbb{R}$ be such that $h \leq f$. If $S(h, \alpha) \cap D = \emptyset$ for some $\alpha \in \mathbb{R}$, then $\alpha < f^*$.*

Proof. Suppose that $S(h, \alpha) \cap D = \emptyset$ and $\alpha \geq f^*$. Then $f^* \in S(f, \alpha) \cap D$. Let h be such that $h \leq f$. Then $S(f, \alpha) \cap D \subset S(h, \alpha) \cap D$, by Lemma 4, and, consequently, $S(h, \alpha) \cap D \neq \emptyset$. We obtain a contradiction. ∎

Lemma 7. *Let function $h : \mathbb{R}^n \to \mathbb{R}$ be such that $h \leq f$. If $S'(h, \alpha) \cap D = \emptyset$ for some $\alpha \in \mathbb{R}$, then $\alpha \leq f^*$.*

Proof. Suppose that $S'(h, \alpha) \cap D = \emptyset$ and $\alpha > f^*$. Let $\beta \in \mathbb{R}$ be such that $\alpha > \beta > f^*$. Then

$$S'(h, \alpha) \cap D \supset S(h, \beta) \cap D \supset S(h, f^*) \cap D \supset S(f, f^*) \cap D = M \neq \emptyset,$$

by Lemmas 4 and 5. We obtain a contradiction. ∎

The model f_L of the form (7) satisfies the condition $f_L \leq f$. Therefore, we can use the function f_L in Lemma 6 and Lemma 7 instead of the function h.

Remark 8.

a) If the condition in Step 6.2 is satisfied, then $\alpha_k < f^*$ (level α_k is too low), by Lemma 6. Therefore, we can execute the lower bound update (go to Step 10).

b) Suppose that condition in Step 6.2 is substituted for

$$S'_k := \bigcap_{i \in L_k} S'(f_i, \alpha_k) = \emptyset.$$

If this condition is satisfied, then $\alpha_k \leq f^*$, by Lemma 7. Therefore, we can execute Step 10.

The following lemmas explain why we have to go to Step 10 when the situation described in Step 8.2 occurs. Recall that $z'_k = P_D z_k$ and $q_k = z'_k - z_k$ (see Step 7).

Lemma 9. *If $\alpha_k \geq f^*$ for $k \geq 1$ then*

$$\|x_{k+1} - z\|^2 \leq \|x_k - z\|^2 - \lambda_k(2 - \lambda_k) \|t_k\|^2 - \|q_k\|^2 \tag{8}$$

for all $z \in M$. Furthermore, if $\alpha_k \geq f^*$ for $k = k_1, \ldots, k_2$, then

$$\|x_{k_2+1} - z\|^2 \leq \|x_{k_1} - z\|^2 - \sum_{k=k_1}^{k_2} (\lambda_k(2 - \lambda_k) \|t_k\|^2 + \|q_k\|^2) \tag{9}$$

for all $z \in M$.

Proof. (See [2, Lemma 1 and Corollary 1]). ■

Remark 10. If

$$\sum_{k=k_1}^{k_2} (\lambda_k(2 - \lambda_k) \|t_k\|^2 + \|q_k\|^2) > \|x_{k_1} - z\|^2 - \|x_{k_2+1} - z\|^2, \tag{10}$$

then $\alpha_k < f^*$ for some k, $k_1 \leq k \leq k_2$.

Lemma 11. *Suppose that the sequence* (α_k) *is non-increasing for* $k = k_1, \ldots, k_2$. *If*

$$\sum_{i=k_1}^{k} (\lambda_i(2-\lambda_i)\|t_i\|^2 + \|q_i\|^2) > R^2 - (R - \|z'_k - x_{k_1}\|)^2 \tag{11}$$

for some k, $k_1 \le k \le k_2$, *then* $\alpha_k < f^*$.

Proof. (See [2, Lemma 4]). ■

Lemma 12. *Suppose that the sequence* (α_k) *is non-increasing for* $i = k_1, \ldots, k_2$. *If*

$$\sum_{i=k_1}^{k-1} (\lambda_i(2-\lambda_i)\|t_i\|^2 + \|q_i\|^2) + \|t_k\|^2 > R^2 - (R - \|x_k + t_k - x_{k_1}\|)^2 \tag{12}$$

for some k, $k_1 \le k \le k_2$, *then* $\alpha_k < f^*$.

Proof. Suppose that the assumptions of the lemma are satisfied but $\alpha_k \ge f^*$. By Lemma 9, we obtain

$$\sum_{i=k_1}^{k-1} (\lambda_i(2-\lambda_i)\|t_i\|^2 + \|q_i\|^2) \le \|x_{k_1} - z\|^2 - \|x_k - z\|^2 \tag{13}$$

for all $z \in M$. Suppose that $\lambda_k = 1$ in Step 7. By inequality (8) in Lemma 9, we obtain

$$\|t_k\|^2 \le \|x_k - z\|^2 - \|x_k + t_k - z\|^2 . \tag{14}$$

By inequalities (13) and (14), we have

$$\sum_{i=k_1}^{k-1} (\lambda_i(2-\lambda_i)\|t_i\|^2 + \|q_i\|^2) + \|t_k\|^2 \le \|x_{k_1} - z\|^2 - \|x_k + t_k - z\|^2 \tag{15}$$

On the other hand, by the assumption of the lemma, the inequality $R \geq \|x_{k_1} - z\|$ and the triangle inequality, we obtain

$$\begin{aligned}
&\sum_{i=k_1}^{k-1} (\lambda_i(2-\lambda_i)\,\|t_i\|^2 + \|q_i\|^2) + \|t_k\|^2 \\
&> R^2 - (R - \|x_k + t_k - x_{k_1}\|)^2 \\
&\geq \|x_{k_1} - z\|^2 - (\|x_{k_1} - z\| - \|x_k + t_k - x_{k_1}\|)^2 \\
&\geq \|x_{k_1} - z\|^2 - \|x_k + t_k - z\|^2 .
\end{aligned}$$

which is a contradiction to inequality (15). ■

Remark 13. The first condition in Step 8.2 corresponds to the condition in Lemma 11 and the second condition in Step 8.2 corresponds to the condition in Lemma 12. Hence, we have to go to Step 10 when one of the inequalities in Step 8.2 is satisfied.

3. Convergence analysis

In this section we show that any sequence generated by Iterative Sheme 2 has a limit point in the solution set M. The idea of the proof of the convergence comes from [2]. Suppose that Iterative Scheme 2 does not terminate.

Denote $\alpha_k \downarrow \alpha$ for a non-increasing real sequence (α_k) converging to α.

Lemma 14. *Suppose $\alpha_k \downarrow \alpha$ for $k \geq k_1$. Then $\alpha \geq f^*$ if and only if*

$$\sum_{i=k_1}^{k} (\lambda_i(2-\lambda_i)\,\|t_i\|^2 + \|q_i\|^2) \leq R^2 \quad \text{for all} \quad k \geq k_1.$$

Proof. ($\Longrightarrow$) The implication follows from Lemma 11.
($\Longleftarrow$) Suppose that $\sum_{i=k_1}^{k} (\lambda_i(2-\lambda_i)\,\|t_i\|^2 + \|q_i\|^2) \leq R^2$ for all $k \geq k_1$. Then $\|t_k\| \to 0$ and, consequently,

$$\frac{f(x_k) - \alpha_k}{\|g_k\|} \to 0,$$

by Remark 3 c). The function f is locally Lipschitz continuous and the sequence (x_k) is bounded. Therefore, the sequence $\|g_k\|$ is bounded. Hence, $f(x_k) - \alpha_k \to 0$, and, consequently, $f(x_k) \to \alpha$. Hence, $\alpha \geq f^*$. ■

Lemma 15. *Suppose $\alpha_k \downarrow \alpha$ for $k \geq k_1$. If $\alpha \geq f^*$, then $f(x_k) \to \alpha$ and each accumulation point x of the sequence (x_k) belongs to $S(f, \alpha)$.*

Proof. Suppose $\alpha_k \downarrow \alpha \geq f^*$ for $k \geq k_1$. Then

$$\sum_{i=k_1}^{k} (\lambda_i(2 - \lambda_i) \|t_i\|^2 + \|q_i\|^2) \leq R^2$$

for all $k \geq k_1$, by Lemma 14. Furthermore, $f(x_k) \to \alpha$ (see the proof of Lemma 14). Let $\widetilde{x}$ be an accumulation point of the sequence (x_k). Such a point exists because the sequence (x_k) is bounded. Since $x_k \in D$ for all k and set D is closed, therefore $\widetilde{x} \in D$. Now, from the continuity of f, we have $f(\widetilde{x}) = \alpha$ and $\widetilde{x} \in S(f, \alpha)$. ■

Denote $\triangle_k = \overline{\alpha}_k - \underline{\alpha}_k$.

Theorem 16. *The sequences $(\overline{\alpha}_k)$, $(\underline{\alpha}_k)$, (α_k) converge to f^*.*

Proof. If Step 10 is executed in the kth iteration then $\underline{\alpha}_{k+1} = \alpha_k$ and, consequently

$$\begin{aligned} \triangle_{k+1} &= \overline{\alpha}_{k+1} - \underline{\alpha}_{k+1} \\ &= \overline{\alpha}_k - \alpha_k \\ &= \overline{\alpha}_k - (1 - \nu_k)\overline{\alpha}_k - \nu_k \underline{\alpha}_k \\ &= \nu_k \triangle_k. \end{aligned}$$

Hence, if Step 10 is executed infinitely many times, then $\triangle_k \to 0$ since $\nu_k \leq \overline{\nu} < 1$. Consequently, the sequences $(\overline{\alpha}_k), (\underline{\alpha}_k), (\alpha_k)$ converge to f^*.

Now suppose that k_1 is the last iteration in which Step 10 is executed. Then $\underline{\alpha}_k$ is constant for $k > k_1$ and $(\alpha_k)_{k>k_1}$ is a non-increasing sequence. Let $\alpha = \lim_k \alpha_k$. By Lemma 14, $\alpha \geq f^*$. Otherwise the first condition in Step 8.2 is satisfied and Step 10 would be executed for some $k > k_1$. Since $f(x_k) \geq \overline{\alpha}_k$ and $\nu_k \geq \underline{\nu}$, we have

$$\begin{aligned} \alpha_k &= (1 - \nu_k)\overline{\alpha}_k + \nu_k \underline{\alpha}_k \\ &\leq (1 - \nu_k) f(x_k) + \nu_k \underline{\alpha}_k \\ &\leq (1 - \underline{\nu}) f(x_k) + \underline{\nu}\, \underline{\alpha}_k. \end{aligned}$$

By Lemma 15, we obtain

$$(1-\underline{\nu})f(x_k)+\underline{\nu}\,\underline{\alpha}_k \to (1-\underline{\nu})\alpha+\underline{\nu}\,\underline{\alpha}_{k_1+1},$$

since $\underline{\alpha}_k$ is constant for $k>k_1$. Furthermore,

$$(1-\underline{\nu})\alpha+\underline{\nu}\,\underline{\alpha}_{k_1+1}\le\alpha,$$

since $\underline{\nu}\in(0,1)$ and $\underline{\alpha}_{k_1+1}\le f^*\le\alpha$. Consequently, we obtain

$$\alpha\leftarrow\alpha_k\le(1-\underline{\nu})f(x_k)+\underline{\nu}\,\underline{\alpha}_k\to(1-\underline{\nu})\alpha+\underline{\nu}\,\underline{\alpha}_{k_1+1}\le\alpha.$$

Therefore, we have $(1-\underline{\nu})\alpha+\underline{\nu}\,\underline{\alpha}_{k_1+1}=\alpha$, and, consequently, $\underline{\alpha}_k=\alpha$ for $k>k_1$, since $\underline{\nu}>0$ and $\underline{\alpha}_k$ is constant for $k>k_1$. Hence, $f^*\ge\underline{\alpha}_k=\alpha\ge f^*$ for $k>k_1$, and, consequently, $\underline{\alpha}_k=\alpha=f^*$ for $k>k_1$.

Since $\nu_k\ge\underline{\nu}$,

$$\alpha_k=(1-\nu_k)\overline{\alpha}_k+\nu_k\underline{\alpha}_k\le(1-\underline{\nu})\overline{\alpha}_k+\underline{\nu}\,\underline{\alpha}_k.$$

Therefore, we obtain

$$\overline{\alpha}_k=\frac{\alpha_k-\nu_k\underline{\alpha}_k}{1-\nu_k}\ge\frac{\alpha_k-\underline{\nu}\,\underline{\alpha}_k}{1-\underline{\nu}},$$

since $\underline{\nu}<1$. Moreover,

$$\frac{\alpha_k-\underline{\nu}\,\underline{\alpha}_k}{1-\underline{\nu}}\to\alpha,$$

since $\alpha_k\to\alpha$ and $\underline{\alpha}_k=\alpha$ for $k>k_1$. Of course, $f(x_k)\ge\overline{\alpha}_k$ and $f(x_k)\to\alpha$, by Lemma 15. Hence,

$$\alpha\leftarrow f(x_k)\ge\overline{\alpha}_k\ge\frac{\alpha_k-\underline{\nu}\,\underline{\alpha}_k}{1-\underline{\nu}}\to\alpha,$$

consequently, $\overline{\alpha}_k\to\alpha=f^*$. Therefore, $\alpha_k\to\alpha=f^*$. ■

Theorem 17. *Each accumulation point $\overline{x}$ of the sequence $(\overline{x}_k)$ belongs to M.*

Proof. By Theorem 16, $f(\overline{x}_k) = \overline{\alpha}_k \to f^*$. Moreover, the sequence $(\overline{x}_k)$ is bounded. Let $\overline{x}$ be an accumulation point of the sequence $(\overline{x}_k)$. Since $\overline{x}_k \in D$ for all k and set D is closed, therefore $\overline{x} \in D$. From the continuity of f, we have $f(\overline{x}) = f^*$ and $\overline{x} \in S(f, f^*) = M$. ∎

4. Efficiency

The idea of the efficiency estimate comes from [6]. The efficiency of the method is the number of objective evaluations (function and subgradient calculations) sufficient to obtain an ε-optimal solution.

All considerations in this Section deal with Iterative Scheme 2. We assume that $\varepsilon > 0$ in Step 0. By Theorem 16, the stopping criterion $\overline{\alpha}_k - \underline{\alpha}_k \leq \varepsilon$ is satisfied for some $k \in \mathbb{N}$ ($\overline{x}_k$ is an ε-optimal solution) and Iterative Scheme 2 generates finite sequence of iterations.

We denote:

- p – the final value of k,
- l' – the final value of l,
- $m = p - l'$ – the number of objective evaluations,
- k_l – the iteration at which lth execution of Step 10 occurs, $l = 1, \ldots, l'$,
- $k_0 = 0$, $k_{l'+1} = p$,
- $\delta_l = \triangle_{k_l} = \overline{\alpha}_{k_l} - \underline{\alpha}_{k_l}$, $l = 1, \ldots, l' + 1$,
- $j_l = k_l - k_{l-1} - 1$, $l = 2, \ldots, l' + 1$, $j_1 = k_1$.

Lemma 18. *For $l = 1, \ldots, l'$ we have*

$$\delta_{l+1} \leq \triangle_{k_l+1} \leq \overline{\nu}\delta_l.$$

Proof. Recall that $\triangle_k = \overline{\alpha}_k - \underline{\alpha}_k$ is nonincreasing for $k \leq p$. For $l = 1, \ldots, l'$ and for $k_l < k \leq k_{l+1}$ we have the inequality

$$\delta_{l+1} = \triangle_{k_{l+1}} \leq \triangle_k \leq \triangle_{k_l+1}.$$

Since $\overline{\alpha}_{k_l+1} = \overline{\alpha}_{k_l}$, $\underline{\alpha}_{k_l+1} = \alpha_{k_l}$ (Step 10.1) and $\nu_k \leq \overline{\nu}$, hence

$$\begin{aligned} \triangle_{k_l+1} &= \overline{\alpha}_{k_l+1} - \underline{\alpha}_{k_l+1} \\ &= \overline{\alpha}_{k_l} - \alpha_{k_l} \\ &= \nu_k(\overline{\alpha}_{k_l} - \underline{\alpha}_{k_l}) \\ &\leq \overline{\nu}(\overline{\alpha}_{k_l} - \underline{\alpha}_{k_l}) = \overline{\nu}\delta_l. \end{aligned}$$

∎

Denote $\lceil \gamma \rceil = \min\{n \in \mathbb{N} : n \geq \gamma\}$.

Theorem 19. *Suppose $l' \geq 1$. Then*

$$l' \leq \left\lceil \frac{\log \frac{\varepsilon}{\triangle_1}}{\log \overline{\nu}} \right\rceil .$$

Proof. By Lemma 18, we obtain

$$\triangle_{k_l+1} \leq \overline{\nu}\delta_l \leq \overline{\nu}^2\delta_{l-1} \leq \ldots \leq \overline{\nu}^l\delta_1 = \overline{\nu}^l\triangle_{k_1}$$

for $l = 1, \ldots, l'$. Furthermore, $\triangle_{k_1} \leq \triangle_1$ since $k_1 \geq 1$. Hence,

$$\triangle_{k_l+1} \leq \overline{\nu}^l\triangle_1$$

for $l = 1, \ldots, l'$. From this inequality for $l = l' - 1$ and from inequalities $k_{l'-1} + 1 \leq k_{l'}$ and $\triangle_{k_{l'}} > \varepsilon$ we obtain

$$\overline{\nu}^{l'-1}\triangle_1 \geq \triangle_{k_{l'-1}+1} \geq \triangle_{k_{l'}} > \varepsilon.$$

Hence,

$$l' - 1 < \log_{\overline{\nu}} \frac{\varepsilon}{\triangle_1},$$

since $\overline{\nu} < 1$, and, consequently,

$$l' \leq \left\lceil \frac{\log \frac{\varepsilon}{\triangle_1}}{\log \overline{\nu}} \right\rceil .$$

■

Now we estimate the number of the objective evaluations, which is enough to obtain an ε-optimal solution.

Remark 20. The number of the objective evaluations is equal to $\sum_{l=1}^{l'+1} j_l$. Indeed,

$$\begin{aligned}\sum_{l=1}^{l'+1} j_l &= k_1 + (k_2 - k_1 - 1) + \ldots + (k_{l'} - k_{l'-1} - 1) + (k_{l'+1} - k_{l'} - 1) \\ &= k_{l'+1} - l' = p - l' = m.\end{aligned}$$

Lemma 21. *For $l' \geq 1$ and $l = 2, \ldots, l'+1$ we have*

$$R^2 \geq \underline{\lambda}(2-\overline{\lambda})j_l \left(\frac{\underline{\nu}\triangle_{k_l-1}}{L}\right)^2, \tag{16}$$

where L is a Lipschitz constant of the function f on D and $R \geq d(x_1, M)$.

Proof. Let $l \in \{2, \ldots, l'+1\}$. For k such that $k_{l-1}+1 \leq k \leq k_l - 1$ the inequalities in Step 8.2 are not satisfied and we have $\sum_{k=k_{l-1}+1}^{k_l-1} \left(\lambda_k(2-\lambda_k)\|t_k\|^2 + \|q_k\|^2\right) \leq R^2$. Therefore, we obtain for $l = 2, \ldots, l'$,

$$\begin{aligned} R^2 &\geq \sum_{k=k_{l-1}+1}^{k_l-1} \left(\lambda_k(2-\lambda_k)\|t_k\|^2 + \|q_k\|^2\right) \\ &\geq \sum_{k=k_{l-1}+1}^{k_l-1} \lambda_k(2-\lambda_k)\|t_k\|^2 \\ &\geq \underline{\lambda}(2-\overline{\lambda}) \sum_{k=k_{l-1}+1}^{k_l-1} \left(\frac{f(x_k)-\alpha_k}{\|g_k\|}\right)^2 \\ &\geq \underline{\lambda}(2-\overline{\lambda}) \sum_{k=k_{l-1}+1}^{k_l-1} \left(\frac{\underline{\nu}\triangle_k}{L}\right)^2 \\ &\geq \underline{\lambda}(2-\overline{\lambda}) \sum_{k=k_{l-1}+1}^{k_l-1} \left(\frac{\underline{\nu}\triangle_{k_l-1}}{L}\right)^2 \\ &= \underline{\lambda}(2-\overline{\lambda}) \left(\frac{\underline{\nu}\triangle_{k_l-1}}{L}\right)^2 j_l, \end{aligned}$$

where the third inequality stems from $\underline{\lambda} \leq \lambda_k \leq \overline{\lambda}$ and Remark 3 c), the fourth from

$$f(x_k) - \alpha_k \geq \overline{\alpha}_k - \alpha_k = \nu_k(\overline{\alpha}_k - \underline{\alpha}_k) \geq \underline{\nu}\triangle_k$$

and $\|g_k\| \leq L$, the fifth from the inequality $\triangle_k \geq \triangle_{k_l-1}$ for $k \leq k_l - 1$, and the final equality from

$$k_l - 1 - (k_{l-1}+1) + 1 = k_l - k_{l-1} - 1 = j_l.$$

■

Remark 22.

a) If $l' \geq 1$ and $l = 1$, then, similarly as in proof of Lemma 21, one can show that

$$R^2 \geq \underline{\lambda}(2-\overline{\lambda})(j_1 - 1)\left(\frac{\underline{\nu}\triangle_{k_1-1}}{L}\right)^2, \tag{17}$$

since

$$k_1 - 1 - (k_0 + 1) + 1 = k_1 - 1 = j_1 - 1.$$

b) If $l' = 0$ then $m = p = k_1 = j_1$ and, for $m > 1$, similarly as in proof of Lemma 21, one can show that

$$R^2 \geq \underline{\lambda}(2-\overline{\lambda})(m - 1)\left(\frac{\underline{\nu}\triangle_{m-1}}{L}\right)^2.$$

Since $\triangle_{m-1} > \varepsilon$, the number of the objective evaluations fulfills the inequality

$$m \leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\varepsilon}\right)^2 + 1.$$

Theorem 23. *If $l' \geq 1$, then*

$$m \leq \frac{1}{\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2(1-\overline{\nu}^2)}\left(\frac{RL}{\varepsilon}\right)^2 + 1, \tag{18}$$

where L is a Lipschitz constant of f on the set D and $R \geq d(x_1, M)$.

Proof. From Remark 20, we have $m = \sum_{l=1}^{l'+1} j_l$.
Now we estimate j_l for $l = 1, \ldots, l'+1$. For $l = 1, \ldots, l'$, we obtain

$$\begin{aligned} \triangle_{k_l-1} \ &\geq\ \triangle_{k_l} = \delta_l \\ &\geq\ \overline{\nu}^{-1}\delta_{l+1} \geq \ldots \geq \overline{\nu}^{(l-l')}\delta_{l'}, \end{aligned} \tag{19}$$

where the inequalities stems from Lemma 18. From Lemma 21 and from the above inequalities, we obtain

$$
\begin{aligned}
j_l &\le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\triangle_{k_l-1}}\right)^2 \\
&\le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\delta_{l'}}\right)^2 \overline{\nu}^{2(l'-l)}
\end{aligned} \tag{20}
$$

for $l = 2, \ldots, l'$. For $l = l' + 1$, we have

$$\triangle_{k_{l'+1}-1} = \triangle_{p-1} > \varepsilon.$$

From Lemma 21 for $l = l' + 1$ and from the above inequality, we obtain

$$
\begin{aligned}
j_{l'+1} &\le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\triangle_{k_{l'+1}-1}}\right)^2 \\
&< \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\varepsilon}\right)^2 .
\end{aligned} \tag{21}
$$

From Remark 22 and inequality 19, we obtain

$$
\begin{aligned}
j_1 - 1 &\le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\triangle_{k_1-1}}\right)^2 \\
&\le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\delta_{l'}}\right)^2 \overline{\nu}^{2(l'-1)}.
\end{aligned} \tag{22}
$$

Now we estimate the number of the objective evaluations. At first, we consider the case when $p > k_{l'} + 1$. Then,

$$\delta_{l'} \ge \overline{\nu}^{-1}\triangle_{k_{l'}+1} > \overline{\nu}^{-1}\varepsilon, \tag{23}$$

where we obtain the first inequality similarly as in the proof of Lemma 18. From inequalities (20) and (23), we obtain

$$j_l \le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\varepsilon}\right)^2 \overline{\nu}^{2(l'-l+1)}$$

for $l = 2, \ldots, l'$. From inequalities (22) and (23), and from $\overline{\nu} \in (0, 1)$, we obtain

$$j_1 \le \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1}\left(\frac{RL}{\varepsilon}\right)^2 \overline{\nu}^{2l'} + 1. \tag{24}$$

Since

$$\sum_{l=1}^{l'+1} \overline{\nu}^{2(l'-l+1)} = \sum_{i=0}^{l'} \overline{\nu}^{2i} \leq \sum_{i=0}^{\infty} \overline{\nu}^{2i} = \frac{1}{1-\overline{\nu}^2},$$

then, consequently, we obtain

$$\begin{aligned} m &= \sum_{l=1}^{l'+1} j_l \leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \sum_{l=1}^{l'+1} \overline{\nu}^{2(l'-l+1)} + 1 \\ &\leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \frac{1}{1-\overline{\nu}^2} + 1. \end{aligned}$$

Now we consider the case when $p = k_{l'} + 1$. Then,

$$\delta_{l'} > \varepsilon \tag{25}$$

and $j_{l'+1} = 0$. Similarly as above, we obtain

$$j_l \leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \overline{\nu}^{2(l'-l)}$$

for $l = 2, \ldots, l'$ and

$$j_1 \leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \overline{\nu}^{2(l'-1)} + 1.$$

Since

$$\sum_{l=1}^{l'} \overline{\nu}^{2(l'-l)} = \sum_{i=0}^{l'-1} \overline{\nu}^{2i} \leq \sum_{i=0}^{\infty} \overline{\nu}^{2i} = \frac{1}{1-\overline{\nu}^2},$$

then, consequently, we obtain

$$\begin{aligned} m &= \sum_{l=1}^{l'} j_l \leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \sum_{l=1}^{l'} \overline{\nu}^{2(l'-l)} + 1 \\ &\leq \left(\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2\right)^{-1} \left(\frac{RL}{\varepsilon}\right)^2 \frac{1}{1-\overline{\nu}^2} + 1. \end{aligned}$$

■

Corollary 24. *If* $\triangle_1 \geq \varepsilon > 0$, *then Iterative Scheme* 2 *requires at most*

$$m(\varepsilon) = \left\lceil \frac{1}{\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2(1-\overline{\nu}^2)} \left(\frac{RL}{\varepsilon}\right)^2 \right\rceil + 1$$

objective evaluations and at most

$$k(\varepsilon) = m(\varepsilon) + \left\lceil \frac{\log \frac{\varepsilon}{\triangle_1}}{\log \overline{\nu}} \right\rceil$$

iterations to obtain an ε*-optimal solution, where* L *is a Lipschitz constant of the function* f *on the set* D, *whereas* $R \geq d(x_1, M)$.

Proof. Suppose that $l' \geq 1$. Then, $m \leq m(\varepsilon)$ by Theorem 23 and

$$l' \leq l(\varepsilon) = \left\lceil \frac{\log \frac{\varepsilon}{\triangle_1}}{\log \overline{\nu}} \right\rceil$$

by Theorem 19. Consequently,

$$p = m + l' \leq m(\varepsilon) + l(\varepsilon) = k(\varepsilon).$$

Suppose now that $l' = 0$. Then, $m = p = k_1 = j_1$. If $p = 1$, then $\triangle_p = \triangle_1 < \varepsilon$. We obtain a contradiction with assumption $\triangle_1 \geq \varepsilon$. If $p > 1$, then

$$\begin{aligned} m &\leq \frac{1}{\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2} \left(\frac{RL}{\varepsilon}\right)^2 + 1 \\ &\leq \frac{1}{\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2(1-\overline{\nu}^2)} \left(\frac{RL}{\varepsilon}\right)^2 + 1 \\ &\leq \left\lceil \frac{1}{\underline{\lambda}(2-\overline{\lambda})\underline{\nu}^2(1-\overline{\nu}^2)} \left(\frac{RL}{\varepsilon}\right)^2 \right\rceil + 1 \end{aligned}$$

by Remark 22 b). ■

Remark 25. The result obtained in Corollary 24 is a generalization of the results presented in [2, 6], where $\nu_k = \nu$ for $k = 1, 2, \ldots$
If $\nu_k = \nu \in (0, 1)$ for $k \geq 1$ in Iterative Scheme 2, then
$m(\varepsilon) = \left\lceil \frac{1}{\underline{\lambda}(2-\overline{\lambda})\nu^2(1-\nu^2)} \left(\frac{RL}{\varepsilon}\right)^2 \right\rceil + 1$ and $k(\varepsilon) = m(\varepsilon) + \left\lceil \frac{\log \frac{\varepsilon}{\triangle_1}}{\log \nu} \right\rceil$.

References

[1] A. Cegielski, *Relaxation Methods in Convex Optimization Problems,* Higher College of Engineering, Series Monographies, No. **67** (1993), Zielona Góra, Poland (Polish).

[2] A. Cegielski, *A method of projection onto an acute cone with level control in convex minimization*, Mathematical Programming **85** (1999), 469–490.

[3] A. Cegielski and R. Dylewski, *Selection strategies in projection methods for convex minimization problems*, Discuss. Math. Differential Inclusions, Control and Optimization **22** (2002), 97–123.

[4] A. Cegielski and R. Dylewski, *Residual selection in a projection method for convex minimization problems*, Optimization **52** (2003), 211–220.

[5] S. Kim, H. Ahn and S.-C. Cho, *Variable target value subgradient method*, Mathematical Programming **49** (1991), 359–369.

[6] K.C. Kiwiel, *The efficiency of subgradient projection methods for convex optimization, part I: General level methods*, SIAM J. Control and Optimization **34** (1996), 660–676.

[7] C. Lemaréchal, A.S. Nemirovskii and Yu.E. Nesterov, *New variants of bundle methods*, Mathematical Programming **69** (1995), 111–147.

[8] B.T. Polyak, *Minimization of unsmooth functionals*, Zh. Vychisl. Mat. i Mat. Fiz. **9** (1969), 509–521.

Received 10 May 2009