

Approximation of Solutions to Constrained Convex Minimization Problem in Hilbert Spaces

Yekini Shehu

Received: 25 September 2013 / Accepted: 23 March 2014 / Published online: 17 August 2014
© Vietnam Academy of Science and Technology (VAST) and Springer Science+Business Media Singapore 2014

Abstract The idea of this paper is to perturb Mann iteration scheme and obtain a strong convergence result for approximation of solutions to constrained convex minimization problem in a real Hilbert space. Furthermore, we give computational analysis of our iterative scheme.

Keywords Gradient projection method · Constrained convex minimization · Strong convergence

Mathematics Subject Classification (2010) 47H09 · 47J25

1 Introduction

Let H be a real Hilbert space and C a nonempty, closed, and convex subset of H . Let $f : H \rightarrow \mathbb{R}$ be a convex and continuously Fréchet differentiable functional. Consider the following constrained convex minimization problem:

$$\text{minimize}\{f(x) : x \in C\}. \quad (1)$$

The gradient projection method (for short, GPM) generates a sequence $\{x_n\}$ using the following recursive formula:

$$x_{n+1} = P_C(x_n - \lambda \nabla f(x_n)) \quad \forall n \geq 1, \quad (2)$$

or more generally,

$$x_{n+1} = P_C(x_n - \lambda_n \nabla f(x_n)) \quad \forall n \geq 1, \quad (3)$$

where in both (2) and (3), the initial guess x_0 is taken from C arbitrarily, and the parameters, λ or λ_n , are positive real numbers known as stepsize.

Y. Shehu (✉)
Department of Mathematics, University of Nigeria, Nsukka, Nigeria
e-mail: deltanougt2006@yahoo.com; yekini.shehu@unn.edu.ng

The gradient projection (or projected-gradient) algorithm is a powerful tool for solving constrained convex optimization problems and has been extensively studied (see [5, 6, 9, 11, 15, 17–23] and the references therein). It has been recently applied to solve split feasibility problems which find applications in image reconstructions and the intensity-modulated radiation therapy (see [3, 4, 14, 25]).

The convergence of algorithms (2) and (3) depends on the behavior of the gradient ∇f . The gradient projection method (3) has been considered with several stepsize rules:

- Constant stepsize, where for some $\lambda > 0$, we have $\lambda_n = \lambda$ for all n .
- Diminishing stepsize, where $\lambda_n \rightarrow 0$ and $\sum_{n=1}^{\infty} \lambda_n = \infty$.
- Polyak’s stepsize, where $\lambda_n = \frac{f(x_n) - f^*}{\|\nabla f(x_n)\|^2}$, where f^* is the optimal value of (1).
- Modified Polyak’s stepsize, where $\lambda_n = \frac{f(x_n) - \hat{f}_n}{\|\nabla f(x_n)\|^2}$ and $\hat{f}_n = \min_{0 \leq j \leq n} f(x_j) - \delta$ for some scalar $\delta > 0$.

The constant stepsize rule is suitable when we are interested in finding an approximate solution to problem (1). Diminishing stepsize rule is an off-line rule and is typically used with $\lambda_n = \frac{c}{n+1}$ or $\frac{c}{\sqrt{n+1}}$ for some $c > 0$. The constant and the diminishing stepsizes are also well suited for some distributed implementations of the method.

As a matter of fact, it is known (see [11]) that if ∇f is α -strongly monotone and L -Lipschitzian with constants $\alpha, L > 0$, then the operator

$$T := P_C(I - \lambda \nabla f) \tag{4}$$

is a contraction; hence, the sequence $\{x_n\}$ defined by algorithm (2) converges in norm to the unique solution of the minimization problem (1). More generally, if the sequence $\{\lambda_n\}$ is chosen to satisfy the property

$$0 < \liminf \lambda_n \leq \limsup \lambda_n < \frac{2\alpha}{L^2},$$

then the sequence $\{x_n\}$ defined by algorithm (3) converges in norm to the unique minimizer of (1). However, if the gradient ∇f fails to be strongly monotone, the operator T defined by (4) would fail to be contractive; consequently, the sequence $\{x_n\}$ generated by algorithm (3) may fail to converge strongly (see [22, Sect. 5]). If ∇f is Lipschitzian, then algorithms (2) and (3) can still converge in the weak topology under certain conditions.

Recently, Xu [22] gave an alternative operator-oriented approach to algorithm 3, namely an averaged mapping approach. He gave his averaged mapping approach to the gradient projection algorithm (3) and the relaxed gradient projection algorithm. Moreover, he constructed a counterexample which shows that algorithm (2) does not converge in norm in an infinite-dimensional space and also presented two modifications of gradient projection algorithms which are shown to have strong convergence. Further, he regularized the minimization problem (1) to devise an iterative scheme that generates a sequence converging in norm to the minimum-norm solution of (1) in the consistent case.

Very recently, motivated by the work of Xu [22], Ceng et al. [6] proposed implicit and explicit iterative schemes for finding the approximate minimizer of a constrained convex minimization problem and prove that the sequences generated by their schemes converge strongly to a solution of the constrained convex minimization problem. Such a solution is also a solution of a variational inequality defined over the set of fixed points of a nonexpansive mapping.

Motivated by the works of Xu [22] and Ceng et al. [6], we prove a strong convergence theorem for finding the approximate minimizer of a constrained convex minimization problem in a real Hilbert space. Furthermore, we give computational analysis of our result. Our result complements the results of Xu [22] and several other works in this direction.

2 Preliminaries

Definition 1 A mapping $T : C \rightarrow C$ is said to be *nonexpansive* if

$$\|Tx - Ty\| \leq \|x - y\| \quad \forall x, y \in C.$$

Construction of fixed points of nonexpansive mappings is an important subject in non-linear mapping theory and its applications, in particular, in image recovery and signal processing (see, for example, [3, 4, 14, 25]). For the past 40 years or so, the approximation of fixed points of nonexpansive mappings and fixed points of some of their generalizations and approximation of zeros of accretive-type operators have been a flourishing area of research for many mathematicians. For example, the reader can consult the recent monographs of Berinde [2] and Chidume [7].

For any point $u \in H$, there exists a unique point $P_C u \in C$ such that

$$\|u - P_C u\| \leq \|u - y\| \quad \forall y \in C.$$

P_C is called the *metric projection* of H onto C . We know that P_C is a nonexpansive mapping of H onto C . It is also known that P_C satisfies

$$\langle x - y, P_C x - P_C y \rangle \geq \|P_C x - P_C y\|^2$$

for all $x, y \in H$. Furthermore, $P_C x$ is characterized by the properties $P_C x \in C$ and

$$\langle x - P_C x, P_C x - y \rangle \geq 0$$

for all $y \in C$.

Definition 2 A mapping $T : H \rightarrow H$ is said to be *firmly nonexpansive* if and only if $2T - I$ is nonexpansive, or equivalently

$$\langle x - y, Tx - Ty \rangle \geq \|Tx - Ty\|^2 \quad \forall x, y \in H.$$

Alternatively, T is firmly nonexpansive if and only if T can be expressed as

$$T = \frac{1}{2}(I + S),$$

where $S : H \rightarrow H$ is nonexpansive. Projections are firmly nonexpansive.

Definition 3 A mapping $T : H \rightarrow H$ is said to be an *averaged mapping* if and only if it can be written as the average of the identity mapping I and a nonexpansive mapping; that is

$$T = (1 - \alpha)I + \alpha S, \tag{5}$$

where $\alpha \in (0, 1)$ and $S : H \rightarrow H$ is nonexpansive. More precisely, when (5) holds, we say that T is α -averaged. Thus, firmly nonexpansive mappings (in particular, projections) are $\frac{1}{2}$ -averaged mappings.

Some properties of averaged mappings are in the following proposition.

Proposition 1 ([4, 8]) *For given operators $S, T, V : H \rightarrow H$.*

- (a) *If $T = (1 - \alpha)S + \alpha V$ for some $\alpha \in (0, 1)$ and if S is averaged and V is nonexpansive, then T is averaged.*
- (b) *T is firmly nonexpansive if and only if the complement $I - T$ is firmly nonexpansive.*
- (c) *If $T = (1 - \alpha)S + \alpha V$ for some $\alpha \in (0, 1)$ and if S is firmly nonexpansive and V is nonexpansive, then T is averaged.*
- (d) *The composite of finitely many averaged mappings is averaged. That is, if each of the mappings $\{T_i\}_{i=1}^N$ is averaged, then so is the composite $T_1 \dots T_N$. In particular, if T_1 is α_1 -averaged and T_2 is α_2 -averaged, where $\alpha_1, \alpha_2 \in (0, 1)$, then the composite $T_1 T_2$ is α -averaged, where $\alpha = \alpha_1 + \alpha_2 - \alpha_1 \alpha_2$.*

Definition 4 A nonlinear operator T whose domain $D(T) \subset H$ and range $R(T) \subset H$ is said to be

- (a) *monotone* if

$$\langle x - y, Tx - Ty \rangle \geq 0 \quad \forall x, y \in D(T),$$

- (b) *β -strongly monotone* if there exists $\beta > 0$ such that

$$\langle x - y, Tx - Ty \rangle \geq \beta \|x - y\|^2 \quad \forall x, y \in D(T),$$

- (c) *ν -inverse strongly monotone* (for short, ν -ism) if there exists $\nu > 0$ such that

$$\langle x - y, Tx - Ty \rangle \geq \nu \|Tx - Ty\|^2 \quad \forall x, y \in D(T).$$

It can be easily seen that (i) if T is nonexpansive, then $I - T$ is monotone; (ii) the projection mapping P_C is a 1-ism. The inverse strongly monotone (also referred to as co-coercive) operators have been widely used to solve practical problems in various fields, for instance, in traffic assignment problems; see, for example, [3, 10] and the references therein.

The following proposition gathers some results on the relationship between averaged mappings and inverse strongly monotone operators.

Proposition 2 ([4]) *Let $T : H \rightarrow H$ be an operator from H to itself.*

- (a) *T is nonexpansive if and only if the complement $I - T$ is $\frac{1}{2}$ -ism.*
- (b) *If T is ν -ism, then for $\gamma > 0$, γT is $\frac{\nu}{\gamma}$ -ism.*
- (c) *T is averaged if and only if the complement $I - T$ is ν -ism for some $\nu > 1/2$. Indeed, for $\alpha \in (0, 1)$, T is α -averaged if and only if $I - T$ is $\frac{1}{2\alpha}$ -ism.*

Since the Lipschitz continuity of the gradient of f implies that it is indeed inverse strongly monotone (ism), its complement can be an averaged mapping. Consequently, the GPA can be rewritten as the composite of a projection and an averaged mapping, which is again an averaged mapping. This shows that averaged mappings play an important role in the gradient projection algorithm. Recall that a mapping T is nonexpansive if and only if it is Lipschitz with a Lipschitz constant not more than one and that a mapping is an averaged mapping if and only if it can be expressed as a proper convex combination of the identity mapping and a nonexpansive mapping. An averaged mapping with a fixed point is asymptotically regular and its Picard iterates at each point converge weakly to a fixed point of the mapping. This convergence property is quite helpful.

In the sequel, we shall also make use of the following lemmas.

Lemma 1 *Let H be a real Hilbert space. Then the following inequality holds:*

$$\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle \quad \forall x, y \in H.$$

Lemma 2 *Let H be a real Hilbert space. The following inequality holds:*

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \quad \forall x, y \in H.$$

Lemma 3 (Reich, [16]) *Let K be closed and convex subset of a reflexive real Banach space E with a uniformly Gâteaux differentiable norm. Assume that every weakly compact and convex subset of E has the fixed-point property for nonexpansive mappings. Let $A : K \rightarrow E$ be an accretive mapping which satisfies the range condition $K \subseteq R(I + sA)$ for all $s > 0$. Suppose that $0 \in R(A)$, then for each $x \in K$, the strong limit, $\lim_{t \rightarrow \infty} J_t x$ exists and belongs to $A^{-1}(0)$. If we denote $\lim_{t \rightarrow \infty} J_t x$ by Qx , then $Q : K \rightarrow A^{-1}(0)$ is the unique sunny nonexpansive retraction of K onto $A^{-1}(0)$.*

We remark that in Hilbert spaces, a sunny nonexpansive retraction is a projection mapping.

Lemma 4 (Moore and Nnoli, [12]) *Let $\{\theta_n\}_{n=1}^\infty$ be a sequence of nonnegative real numbers satisfying the following relation:*

$$\theta_{n+1} \leq \theta_n - \alpha_n \Phi(\theta_{n+1}) + \sigma_n, \quad n \geq 1,$$

where (i) $0 < \alpha_n < 1$; (ii) $\sum_{n=1}^\infty \alpha_n = \infty$; (iii) $\Phi : [0, \infty) \rightarrow [0, \infty)$ is a strictly increasing function with $\Phi(0) = 0$. Suppose that $\sigma_n = o(\alpha_n)$ (where $\sigma_n = o(\alpha_n)$ if and only if $\lim_{n \rightarrow \infty} \frac{\sigma_n}{\alpha_n} = 0$), then $\theta_n \rightarrow 0$ as $n \rightarrow \infty$.

We adopt the following notation: $x_n \rightarrow x$ means that $x_n \rightarrow x$ strongly.

3 Main Result

In this section, we modify the gradient projection method so as to have strong convergence. Below, we include such modification. We use constant stepsize λ since we are interested in finding an approximate solution to problem (1).

Theorem 1 *Let C be a nonempty, closed, and convex subset of a real Hilbert space H . Suppose that the minimization problem (1) is consistent and let S denote its solution set. Assume that the gradient ∇f is L -Lipschitzian with constant $L > 0$. For a fixed $u \in C$, let the sequence $\{x_n\}$ be generated iteratively by $x_1 \in C$,*

$$x_{n+1} = (1 - \alpha_n)x_n + \alpha_n P_C(x_n - \lambda \nabla f(x_n)) - \alpha_n \gamma_n(x_n - u), \quad n \geq 1, \tag{6}$$

where $0 < \lambda < \frac{2}{L}$ and $\{\alpha_n\}, \{\gamma_n\}$ are sequences in $(0, 1)$ satisfying the following conditions:

- (i) $\alpha_n(1 + \gamma_n) < 1$,
- (ii) $\alpha_n = o(\gamma_n)$,
- (iii) $\sum_{n=1}^\infty \alpha_n \gamma_n = +\infty$.

Then the sequence $\{x_n\}$ converges strongly to a minimizer \hat{x} of (1), where \hat{x} is the projection of the starting point u onto the solution set of the convex problem being solved.

Proof Observe that $x^* \in C$ solves the minimization problem (1) if and only if x^* solves the fixed-point equation

$$x^* = P_C(I - \lambda \nabla f)x^*,$$

where $\lambda > 0$ is any fixed positive number. Note that the gradient ∇f is L -Lipschitzian with constant $L > 0$ implies that the gradient ∇f is $\frac{1}{L}$ -ism [1], which then implies that $\lambda \nabla f$ is $\frac{\lambda}{L}$ -ism. So by Proposition 2(c), $I - \lambda \nabla f$ is $\frac{\lambda L}{2}$ -averaged. Now since the projection P_C is $\frac{1}{2}$ -averaged, we see from Proposition 2(d) that the composite $P_C(I - \lambda \nabla f)$ is $\frac{2+\lambda L}{4}$ -averaged for $0 < \lambda < \frac{2}{L}$. Therefore, we can write

$$P_C(I - \lambda \nabla f) = \frac{2 - \lambda L}{4}I + \frac{2 + \lambda L}{4}T = (1 - \beta)I + \beta T,$$

where T is nonexpansive and $\beta = \frac{2+\lambda L}{4} \in (0, 1)$. Then, we can rewrite (6) as

$$x_{n+1} = (1 - \theta_n)x_n + \theta_n T x_n - \alpha_n \gamma_n(x_n - u), \tag{7}$$

where $\theta_n = \beta \alpha_n \in (0, 1) \forall n \geq 1$. For any $x^* \in S$, notice that $Tx^* = x^*$. We first show that the sequence $\{x_n\}_{n=1}^\infty$ is bounded. Observe that T is nonexpansive and is equivalent to

$$\langle Tx - Ty, x - y \rangle \leq \|x - y\|^2 - \frac{1}{2}\|(I - T)x - (I - T)y\|^2. \tag{8}$$

Using (7), we have

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &= \|(1 - \alpha_n \gamma_n - \theta_n)(x_n - x^*) + \theta_n(Tx_n - x^*) - \alpha_n \gamma_n(x^* - u)\| \\ &\leq \|(1 - \alpha_n \gamma_n - \theta_n)(x_n - x^*) + \theta_n(Tx_n - x^*)\| + \alpha_n \gamma_n \|u - x^*\|. \end{aligned} \tag{9}$$

Then, using (8), we obtain

$$\begin{aligned} &\|(1 - \alpha_n \gamma_n - \theta_n)(x_n - x^*) + \theta_n(Tx_n - x^*)\|^2 \\ &= (1 - \alpha_n \gamma_n - \theta_n)^2 \|x_n - x^*\|^2 + \theta_n^2 \|Tx_n - x^*\|^2 \\ &\quad + 2(1 - \alpha_n \gamma_n - \theta_n)\theta_n \langle Tx_n - x^*, x_n - x^* \rangle \\ &\leq (1 - \alpha_n \gamma_n - \theta_n)^2 \|x_n - x^*\|^2 + \theta_n^2 \|x_n - x^*\|^2 \\ &\quad + 2(1 - \alpha_n \gamma_n - \theta_n)\theta_n \left[\|x_n - x^*\|^2 - \frac{1}{2}\|x_n - Tx_n\|^2 \right] \\ &= (1 - \alpha_n \gamma_n)^2 \|x_n - x^*\|^2 - (1 - \alpha_n \gamma_n - \theta_n)\theta_n \|x_n - Tx_n\|^2 \\ &\leq (1 - \alpha_n \gamma_n)^2 \|x_n - x^*\|^2, \end{aligned}$$

which implies that

$$\|(1 - \alpha_n \gamma_n - \theta_n)(x_n - x^*) + \theta_n(Tx_n - x^*)\| \leq (1 - \alpha_n \gamma_n)\|x_n - x^*\|. \tag{10}$$

It follows from (9) and (10) that

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq (1 - \alpha_n \gamma_n)\|x_n - x^*\| + \alpha_n \gamma_n \|u - x^*\| \\ &\leq \max\{\|x_n - x^*\|, \|u - x^*\|\} \\ &\vdots \\ &\leq \max\{\|x_1 - x^*\|, \|u - x^*\|\}. \end{aligned}$$

Thus, $\{x_n\}_{n=1}^\infty$ is bounded and so is $\{Tx_n\}$.

Next, we show that the sequence $\{x_n\}_{n=1}^\infty$ converges strongly to \hat{x} . For each $n \geq 1$, let $A := I - T$. Then A is a bounded, continuous, and monotone mapping. Since $\{x_n\}_{n=1}^\infty$ is

bounded, we have that $\{Ax_n\}_{n=1}^\infty$ is bounded. Furthermore, by Theorem 2 of [13], A satisfies the range condition. Observe that if for all $\gamma > 0$, we define

$$A_\gamma : C \rightarrow C \quad \text{by} \quad A_\gamma x = \gamma Ax \quad \forall x \in C,$$

then we easily see that A_γ satisfies the range condition and

$$S = F(T) = A^{-1}(0) = A_\gamma^{-1}(0) = F(J_s^{A_\gamma}) = \{x \in H : J_s^{A_\gamma} x = x\},$$

where $J_s^{A_\gamma}$ is the resolvent of the operator $A_\gamma \forall \gamma > 0$. Observe that

$$\|A_\gamma x_n\| = \gamma \|Ax_n\| \leq \gamma \sup_{x \in B} \|Ax\| \quad \forall n \geq 1,$$

where B is any closed ball containing the sequence $\{x_n\}$. This implies that $\lim_{\gamma \rightarrow 0} \|A_\gamma x_n\| = 0$ for each $n \geq 1$. Furthermore, we obtain from Lemma 3 that $\lim_{s \rightarrow \infty} J_s^{A_\gamma} u = Q^\gamma u$, where Q^γ is the unique projection mapping from $R(I + sA)$ onto $A_\gamma^{-1}(0) = A^{-1}(0) = F(T)$. Thus, from uniqueness of projection mapping, we obtain that if Q is the projection mapping of $R(I + sA)$ onto $A^{-1}(0)$, then $Q^\gamma = Q$ for all $\gamma > 0$. This implies that $Qu = \lim_{s \rightarrow \infty} J_s^{A_\gamma} u$ for all $\gamma > 0$. Let $\hat{x} := Qu$. We show that if

$$\xi_n := \max\{\langle u - \hat{x}, x_n - \hat{x} \rangle, 0\} \quad \forall n \geq 1,$$

then $\lim_{n \rightarrow \infty} \xi_n = 0$. We further observe that since $J_s^{A_\gamma} = (I + sA_\gamma)^{-1}$, then $(I + sA_\gamma)J_s^{A_\gamma} u = u$. This implies that $A_\gamma \circ J_s^{A_\gamma} u = \frac{1}{s}(u - J_s^{A_\gamma} u)$ and thus since A_γ is monotone, we have that

$$\left\langle A_\gamma x_n - \frac{1}{s}(u - J_s^{A_\gamma} u), x_n - J_s^{A_\gamma} u \right\rangle \geq 0 \quad \forall s > 0, \gamma > 0.$$

This implies that for some positive constant $M > 0$,

$$\langle u - J_s^{A_\gamma} u, x_n - J_s^{A_\gamma} u \rangle \leq s \langle A_\gamma x_n, x_n - J_s^{A_\gamma} u \rangle \leq \|A_\gamma x_n\| s M.$$

Thus, $\limsup_{\gamma \rightarrow 0} \langle u - J_s^{A_\gamma} u, x_n - J_s^{A_\gamma} u \rangle \leq 0 \forall n \geq 1$. Therefore, given $\epsilon > 0$, there exists $\delta := \delta(\epsilon) > 0$ such that for all $\gamma \in (0, \delta]$,

$$\langle u - J_s^{A_\gamma} u, x_n - J_s^{A_\gamma} u \rangle < \epsilon.$$

Moreover, we have (in particular, for $\gamma = \delta$) that for some constant $M_0 > 0$,

$$\begin{aligned} \langle u - \hat{x}, x_n - \hat{x} \rangle &= \langle u - \hat{x}, (x_n - \hat{x}) - (x_n - J_s^{A_\delta} u) \rangle + \langle u - J_s^{A_\delta} u, x_n - J_s^{A_\delta} u \rangle \\ &\quad + \langle J_s^{A_\delta} u - \hat{x}, x_n - J_s^{A_\delta} u \rangle \\ &< \langle u - \hat{x}, (x_n - \hat{x}) - (x_n - J_s^{A_\delta} u) \rangle + \|J_s^{A_\delta} u - \hat{x}\| M_0 + \epsilon. \end{aligned} \tag{11}$$

Observe that

$$\lim_{s \rightarrow \infty} \langle u - \hat{x}, (x_n - \hat{x}) - (x_n - J_s^{A_\delta} u) \rangle = 0.$$

Thus, as $s \rightarrow \infty$, we obtain from (11) that $\langle u - \hat{x}, x_n - \hat{x} \rangle \leq \epsilon$, so

$$\limsup_{n \rightarrow \infty} \langle u - \hat{x}, x_n - \hat{x} \rangle \leq \epsilon \tag{12}$$

and since $\epsilon > 0$ is arbitrary, (12) gives

$$\limsup_{n \rightarrow \infty} \langle u - \hat{x}, x_n - \hat{x} \rangle \leq 0$$

from which we can deduce that $\lim_{n \rightarrow \infty} \xi_n = 0$. From (7), we have

$$\begin{aligned} \|x_{n+1} - \hat{x}\|^2 &= \|x_n - \hat{x} - \alpha_n(\beta(x_n - Tx_n) + \gamma_n(x_n - u))\|^2 \\ &\leq \|x_n - \hat{x}\|^2 - 2\alpha_n \langle \beta(x_n - Tx_n) + \gamma_n(x_n - u), x_{n+1} - \hat{x} \rangle \\ &\leq \|x_n - \hat{x}\|^2 - 2\alpha_n \gamma_n \|x_{n+1} - \hat{x}\|^2 + 2\alpha_n \gamma_n \langle (x_{n+1} - x_n) + u - \hat{x}, x_{n+1} - \hat{x} \rangle \\ &\quad - 2\alpha_n \langle (x_n - x_{n+1}) + \beta(I - T)x_{n+1} + x_{n+1} - \beta(I - T)x_{n+1} - x_n \\ &\quad + \beta(I - T)x_n, x_{n+1} - \hat{x} \rangle \\ &\leq \|x_n - \hat{x}\|^2 - 2\alpha_n \gamma_n \|x_{n+1} - \hat{x}\|^2 + (\alpha_n^2 \gamma_n + \alpha_n^2)M^* + 2\alpha_n \gamma_n \langle u - \hat{x}, x_{n+1} - \hat{x} \rangle \\ &\leq \|x_n - \hat{x}\|^2 - 2\alpha_n \gamma_n \|x_{n+1} - \hat{x}\|^2 + \alpha_n^2(\gamma_n + 1)M^* + 2\alpha_n \gamma_n \xi_n \\ &= \|x_n - \hat{x}\|^2 - 2\alpha_n \gamma_n \|x_{n+1} - \hat{x}\|^2 + \delta_n, \end{aligned}$$

where $\delta_n = \alpha_n^2(\gamma_n + 1)M^* + 2\alpha_n \gamma_n \xi_n = o(\alpha_n)$ for some $M^* > 0$. Hence, by Lemma 4, we have that $x_n \rightarrow \hat{x}$ as $n \rightarrow \infty$. This completes the proof. \square

4 Computational Analysis

In this section, we give a computational analysis result using our iterative scheme.

Let H be a real Hilbert space and $C := \{x \in H : \|x\| \leq r\}$ (i.e., a closed ball centered at the origin of radius r) and define $f : C \rightarrow \mathbb{R}$ by $f(x) = \frac{1}{2}\|x\|^2$. Then f is convex and differentiable with $\nabla f(x) = x$. Observe that ∇f is 1-Lipschitzian. Let us consider the problem

$$\min_{x \in C} f(x). \tag{13}$$

Clearly, we see that the optimal solution \hat{x} to the minimization problem (13) is $\hat{x} = 0$.

Now, by our Theorem 1, we can take $\alpha_n = \frac{1}{(n+1)^{1/2}}$, $\gamma_n = \frac{1}{(n+1)^{1/3}}$, $n \geq 1$ and $\lambda = \frac{1}{2}$. By the choice of our function f and λ , we see that $P_C(x - \frac{1}{2}\nabla f(x)) = P_C(x - \frac{1}{2}x) = P_C(\frac{1}{2}x) = \frac{1}{2}x$. Furthermore, our iterative scheme (6) becomes

$$x_{n+1} = \left(1 - \frac{1}{(n+1)^{\frac{1}{2}}}\right)x_n + \frac{x_n}{2(n+1)^{\frac{1}{2}}} - \frac{1}{(n+1)^{\frac{5}{6}}}(x_n - u).$$

5 Conclusions

Since the gradient projection method (GPM) fails, in general, to converge in norm in infinite-dimensional Hilbert spaces, here in our result, we have provided a strongly convergent modification of gradient projection method (GPM).

We note that Theorem 4.1 and Theorem 4.2 of Xu [22] are weak convergence results, while our Theorem 1 is a strong convergence result. Thus, our Theorem 1 improves Theorem 4.1 and Theorem 4.2 of Xu [22]. Furthermore, our iterative scheme in this paper does not involve the “ CQ ” algorithm studied by Xu [22]. Also, our result does not require additional projections which was required in Theorem 5.4 of Xu [22] and Theorem 3.3 of [24] in order to guarantee strong convergence. Our method of proof is different from the methods of proof of Xu [22], Yao et al. [23], Ceng et al. [6], Su and Xu [18], and others.

Acknowledgements The author is very grateful to the Editor and the three anonymous referees for many insightful, detailed, and helpful comments which led to significant improvement of the previous version of the paper.

References

1. Baillon, J.B., Haddad, G.: Quelques proprietes des operateurs angle-bornes et n -cycliquement monotones. *Israel. J. Math.* **26**, 137–150 (1977)
2. Berinde, V.: Iterative approximation of fixed points. *Lecture Notes in Mathematics* 1912. Springer, Berlin (2007)
3. Bertsekas, D.P., Gafni, E.M.: Projection methods for variational inequalities with applications to the traffic assignment problem. *Math. Program. Stud.* **17**, 139–159 (1982)
4. Byrne, C.: Unified treatment of some algorithms in signal processing and image construction. *Inverse Probl.* **20**, 103–120 (2004)
5. Calamai, P.H., More, J.J.: Projected gradient methods for linearly constrained problems. *Math. Program.* **39**, 93–116 (1987)
6. Ceng, L.-C., Ansari, Q.H., Yao, J.-C.: Some iterative methods for finding fixed points and for solving constrained convex minimization problems. *Nonlinear Anal.* **74**, 5286–5302 (2011)
7. Chidume, C.E.: Geometric properties of Banach spaces and nonlinear iterations. *Lecture Notes in Mathematics* 1965. Springer (2009)
8. Combettes, P.L.: Solving monotone inclusion via compositions of nonexpansive averaged operators. *Optim.* **53**, 475–504 (2004)
9. Gafni, E.M., Bertsekas, D.P.: Two metric projection methods for constrained optimization. *SIAM J. Control Optim.* **22**, 936–964 (1984)
10. Han, D., Lo, H.K.: Solving non-additive traffic assignment problems: a descent method for co-coercive variational inequalities. *Eur. J. Oper. Res.* **159**, 529–544 (2004)
11. Levitin, E.S., Polyak, B.T.: Constrained minimization problems. *USSR Comput. Math. Math. Phys.* **6**, 1–50 (1966)
12. Moore, C., Nnoli, B.V.C.: Iterative solution of nonlinear equations involving set-valued uniformly accretive operators. *Comput. Math. Appl.* **42**, 131–140 (2001)
13. Morales, C.H.: Surjectivity theorems for multivalued mappings of accretive type. *Comment. Math. Uni. Carol.* **26**, 397–413 (1985)
14. Podilchuk, C.I., Mammone, R.J.: Image recovery by convex projections using a least-squares constraint. *J. Optim. Soc. Amer.* **A7**, 517–521 (1990)
15. Polyak, B.T.: *Introduction to Optimization*. Optimization Software, New York (1987)
16. Reich, S.: Strong convergence theorems for resolvents of accretive mappings in Banach spaces. *J. Math. Anal. Appl.* **75**, 287–292 (1980)
17. Ruszczynski, A.: *Nonlinear Optimization*. Princeton University Press, New Jersey (2006)
18. Su, M., Xu, H.K.: Remarks on the gradient-projection algorithm. *J. Nonlinear Anal. Optim.* **1**, 35–43 (2010)
19. Wang, C.Y., Xiu, N.H.: Convergence of gradient projection methods for generalized convex minimization. *Comput. Optim. Appl.* **16**, 111–120 (2000)
20. Xiu, N.H., Wang, C.Y., Zhang, J.Z.: Convergence properties of projection and contraction methods for variational inequality problems. *Appl. Math. Optim.* **43**, 147–168 (2001)
21. Xiu, N., Wang, D., Kong, L.: A note on the gradient projection method with exact stepsize rule. *J. Comput. Math.* **25**, 221–230 (2007)
22. Xu, H.K.: Averaged mappings and the gradient-projection algorithm. *J. Optim. Theory Appl.* **150**, 360–378 (2011)
23. Yao, Y., Kang, S.M., Jigang, W., Yang, P.-X.: A Regularized gradient projection method for the minimization problem. *J. Appl. Math.* **2012**(259813), 9 (2012). doi:[10.1155/2012/259813](https://doi.org/10.1155/2012/259813)
24. Yao, Y., Liou, Y.-C., Wen, C.-F.: Variant gradient projection methods for the minimization problems. *Abstr. Appl. Anal.* **2012**(792078), 21 (2012). doi:[10.1155/2012/792078](https://doi.org/10.1155/2012/792078)
25. Youla, D.: On deterministic convergence of iterations of related projection mappings. *J. Vis. Commun. Image Represent* **1**, 12–20 (1990)