國立中山大學應用數學系

碩士論文

Department of Applied Mathematics

National Sun Yat-sen University

Master Thesis

約束極小化之迭代梯度方法

Iterative gradient methods for constrained minimization

研究生：許漢婷

Han-Ting Hsu

指導教授：徐洪坤 博士

Dr. Hong-Kun Xu

姚任之 博士

Dr. Jen-Chih Yao

中華民國 102 年 7 月

July 2013

Iterative gradient methods for constrained minimization

by

Han-Ting Hsu

Advisor: Dr. Hong-Kun Xu, Professor

Dr. Jen-Chih Yao, Professor

Department of Applied Mathematics

National Sun Yat-sen University

Kaohsiung, Taiwan

July, 2013

國立中山大學研究生學位論文審定書

本校應用數學系碩士班

研究生許漢婷（學號：M002040029）所提論文

約束極小化之迭代梯度方法
Iterative gradient methods for constrained minimization

於中華民國 102 年 7 月 8 日經本委員會審查並舉行口試，符合碩士學位論文標準。

學位考試委員簽章：

召集人 林來居 林來居　　　委　員 姚任之 姚任之

委　員 徐洪坤 徐洪坤　　　委　員 黃毅青 黃毅青

委　員 ＿＿＿＿＿＿　　　委　員 ＿＿＿＿＿＿

指導教授(姚任之) 姚任之 (簽名)

指導教授(徐洪坤) 徐洪坤 (簽名)

摘要

這篇論文中，我們探討強凸函數在某非擴張映射之固定點集合上的極小化問題，我們使用正則化的技巧，並引用隱式和顯式兩種迭代方法求解唯一極小點，我們證明了兩種方法的強收斂性，我們的方法推廣了求解約束最佳化問題的梯度投影法。

關鍵字 ：約束極小化、收縮映射、投影、梯度投影法、非擴張映射、固定點、迭代、演算法、收斂性。

Abstract

In this paper we deal with the problem of minimizing a strongly convex objective function over the set of fixed points of a nonexpansive mapping T . This extends the constrained strongly convex minimization problem over a closed convex subset C of a Hilbert space since the projection P_C is nonexpansive. By introducing a parameter $\lambda > 0$, we define a family of contractions T_λ , each of which has a unique fixed point denoted as ξ_λ . We first prove that as $\lambda \to 0$, ξ_λ converges in norm to the unique solution x^* of our minimization problem. We then introduce an iteration algorithm by discretization of the mappings T_λ , and prove strong convergence of this algorithm under appropriate conditions imposed on the sequence of parameters of the algorithm.

Keywords : constrained minimization, contraction, projection, gradient projection method, nonexpansive mapping, fixed point, iteration, algorithm, convergence.

Contents

1 Introduction

Consider the minimization problem

$$\mathrm{Min}_C(\theta) \qquad \min_{x\in C} \theta(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \ ,$$

where C is a closed convex subset of a Hilbert space H , $A\colon H \to H$ is linear bounded, self-adjoint operator and strongly positive, that is,

$$\langle Ax, x\rangle \geqslant \alpha \left\| x\right\|^2 \ \text{ for all } x \in H \ ,$$

where $\alpha > 0$ is some constant.

Since $\mathrm{Min}_C(\theta)$ is a convex minimization problem with a strongly convex objective, it has a unique solution which is the fixed point of the nonlinear mappings

$$x = P_\lambda(x) := P_C\left(I - \lambda\nabla\theta\right)x \tag{1.1}$$

for any $\lambda > 0$. Here P_C is the metric projection from H onto C ; that is,

$$P_C(x) = \arg\min_{y\in H} \|x-y\|, \quad x \in H \ .$$

The gradient projection method for solving $\mathrm{Min}_C(\theta)$ uses the fixed point algorithm that is generated by the mapping P_λ and that generates a sequence $\{x_n\}$ by the iteration process

$$x_{n+1} = P_\lambda x_n := P_C(I - \lambda\nabla\theta)x_n, \quad n \geq 0 \ . \tag{1.2}$$

As the gradient of θ is given by

$$\nabla\theta(x) = Ax - b \ ,$$

the algorithm (1.2) is reduced to the following algorithm

$$x_{n+1} = P_\lambda x_n := P_C(x_n - \lambda(Ax_n - b)), \quad n \geq 0 \ . \tag{1.3}$$

In this paper we consider the case where C is the set of fixed points of some nonexpansive mapping $T\colon H \to H$. Namely,

$$C = Fix\,(T) = \{x \in H \mid Tx = x\} \ .$$

Recall that T is nonexpansive if

$$\|Tx - Ty\| \leqslant \|x - y\| , \quad x, y \in H \ .$$

This setting of C is quite general. In fact, for any closed convex subset C of H , we have $C = Fix\,(P_C)$ and P_C is nonexpansive:

$$\|P_C x - P_C y\| \leqslant \|x - y\| , \quad x, y \in H \ .$$

The minimization problem $\mathrm{Min}_C(\theta)$ under this setting includes, as a matter of fact, many classes of optimization problems such as the hierarchical optimization problems which minimize a function over the set of minimizers of another function, namely, the minimization problems of the form

$$\mathrm{Min}_\Psi(\Phi) \qquad\qquad \min\{\Phi\,(x) \mid x \in \arg\min\,(\Psi)\} \ ,$$

where Φ , Ψ are two convex objective functions defined on H . In the case where Φ and Ψ are continuously differentiable and $\nabla\Psi$ satisfies the Lipschitz continuity condition

$$\|\nabla\Psi\,(x) - \nabla\Psi\,(y)\| \leqslant L_\Psi \,\|x - y\| , \quad x, y \in H \ ,$$

we find that $\mathrm{Min}_\Psi(\Phi)$ can be recovered as $\mathrm{Min}_C\,(\theta)$ with $C = Fix\,(T)$ with T given by

$$T = I - \gamma\nabla\Psi \ ,$$

with $0 < \gamma < \frac{2}{L_\Psi}$. That is,

$$Fix\,(I - \gamma\nabla\Psi) = \{\, x \in H \mid \nabla\Psi\,(x) = 0 \,\} = \arg\min_{x\in H} \Psi\,(x) \ .$$

The study of $\mathrm{Min}_C(\theta)$ in the case where C is the set of a nonexpansive mapping was initiated by Yamada, et al [9]. They considered the mapping T^λ defined by

$$T^\lambda(x) := (I - \lambda A)Tx + \lambda b, \quad x \in H \ , \tag{1.4}$$

where $\lambda \in (0, 1)$. Then proved that if, in addition, $\lambda < \frac{1}{\|A\|}$, then T^λ is a contraction; moreover, if we denote by x_λ the unique fixed point of T^λ , then (x_λ) converges, as $\lambda \to 0$, in norm to the unique solution x^* of $\mathrm{Min}_C(\theta)$.

However, their method does not generalize the gradient projection method which considers the fixed point equation (1.1). This is because (1.4) does not reduce to (1.1) when one takes $T = P_C$. This means that there is some sort of "inconsistency" in their generalization of $\mathrm{Min}_C(\theta)$ to the case where C is the set of a nonexpansive mapping.

The purpose of this paper is to generalize the gradient projection method for $\mathrm{Min}_C(\theta)$ to the case where C is the set of a nonexpansive mapping T . More precisely, instead of considering the mapping T^λ defined by (1.4), we consider the mapping T_λ defined by

$$T_\lambda x := T(\lambda b + (I - \lambda A)x), \quad x \in H \ . \tag{1.5}$$

Note that when $T = P_C$, our mapping T_λ defined by (1.6) is reduced to the mapping P_λ defined by (1.1). Consequently, we generalize the gradient projection method to the case where the set of constraints is the set of fixed points of a nonexpansive mapping.

The structure of this paper is as follows. In the next section we include some basic results which are required in the proof of the main results of this paper. In Section 3, we prove our main results of this paper which are twofold. First we show that if we choose λ in such a way that

$$0 < \lambda < \frac{2\alpha}{\|A\|^2} \ ,$$

then the mapping T_λ is a contraction and hence has a unique fixed point that is denoted by ξ_λ . Our first result is that ξ_λ converges in norm, as $\lambda \to 0$, to the unique solution x^* of $\mathrm{Min}_C(\theta)$. Secondly, we consider an iteration algorithm by discretizing the mapping T_λ . Namely, we define a sequence $\{x_n\}$ by the algorithm

$$x_{n+1} = T_{\lambda_n} x_n = T(\lambda b + (I - \lambda_n A)x_n), \quad n \geq 0 \ . \tag{1.6}$$

Our second main result is to prove the strong convergence to the unique solution x^* of $\mathrm{Min}_C(\theta)$ as $n \to \infty$ under appropriate conditions on the sequence $\{\lambda_n\}$.

2 Preliminaries

Throughout the rest of this paper, it is assumed that H is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, and C be a nonempty, closed and convex subset of H. P_C is the metric projection from H onto C ; that is, for each $x \in H$, $P_C x$ is the unique element of C that satisfies $\|x - P_C x\| = \min\{ \|x - y\| \mid y \in C \}$. We us the notation:

- "$x_n \rightharpoonup x$" denotes weak convergence of (x_n) to x ;
- "$x_n \to x$" denotes strong convergence of (x_n) to x .

Lemma 2.1. *Let H be a Hilbert space and $A\colon H \to H$ a bounded linear operator, there is a constant $\alpha > 0$ such that*

$$\langle Ax, x \rangle \geqslant \alpha \|x\|^2 \text{ for all } x \in H \ .$$

Then A is strongly positive.

Definition 2.1. *$T\colon C \to C$ is said to be a contraction if there exists a constant k , $0 < k < 1$, such that*

$$\|Tx - Ty\| \leqslant k \|x - y\| \text{ for all } x, y \in C \ .$$

Definition 2.2. *$T\colon C \to C$ is said to be nonexpansive if*

$$\|Tx - Ty\| \leqslant \|x - y\| \text{ for all } x, y \in C \ .$$

Lemma 2.2. *Let H be a Hilbert space and C a nonempty, closed and convex subset of H . Assume $x \in H$ and $y \in C$. Then $y = P_C x$ if and only if there holds the inequality*

$$\langle x - y, y - z \rangle \geqslant 0 \quad \text{for all } z \in C.$$

Lemma 2.3. *(Optimality Condition) The first-order optimality condition for x^* be to a solution of $\mathrm{Min}_C(\theta)$ is that*

$$x^* \in Fix\,(T) \ , \langle \nabla\theta(x^*), x - x^* \rangle \geqslant 0 \quad \text{for all } x \in Fix\,(T) \ .$$

Lemma 2.4. *(The Demiclosedness Principle) Let $T\colon C \to C$ be nonexpansive with $Fix\,(T) \neq \emptyset$, where C is a closed convex subset of a real Hilbert space H . If $\{x_n\}$ is sequence such that $x_n \rightharpoonup x$ and $(I - T)x_n \to x$, then $(I - T)x = y$.*

Lemma 2.5. *In a Hilbert space H there holds the inequality*

$$\|x+y\|^2 \leqslant \|x\|^2 + 2\langle y, x+y\rangle \quad \textit{for all } x, y \in H.$$

Lemma 2.6. *The gradient of Ψ is Lipschitz continuous, i.e., there are exists a constant $L_\Psi \geqslant 0$ such that*

$$\|\nabla\Psi(x) - \nabla\Psi(y)\| \leqslant L_\Psi \|x-y\| \ , \ x, y \in H.$$

3 Main Results

We now return to our optimization problem

$$\mathrm{Min}_C(\theta) \qquad \min_{x\in C} \theta\,(x) := \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \ ,$$

where A is a linear bounded self-adjoint operator such that

$$\langle Ax, x\rangle \geqslant \alpha \left\|x\right\|^2 \text{ for all } x \in H \ .$$

This condition guarantees the strong convexity of the objective function

$$\theta\,(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle \ .$$

Consequently, $\mathrm{Min}_C\,(\theta)$ has a unique solution which is denoted x^* .

Now assume $C = Fix\,(T)$, where $T\colon H \to H$ is a nonexpansive mapping. For $\lambda > 0$, let

$$G_\lambda = I - \lambda A \ .$$

Suppose $0 < \lambda < \frac{2\alpha}{\|A\|^2}$ then

Lemma 3.1. *If A is strongly positive, then for $0 < \lambda < \frac{1}{\|A\|}$, we have*

$$\|I - \lambda A\| \leqslant 1 - \lambda\alpha \ .$$

Lemma 3.2. *G_λ is a contraction with contraction coefficient*

$$c_\lambda = \sqrt{1 - \lambda\left(2\alpha - \lambda\left\|A\right\|^2\right)} \in [\ 0, 1\ ) \ .$$

Proof. $\forall x, y \in H$, we compute

$$\begin{aligned}
\left\|G_\lambda x - G_\lambda y\right\|^2 &= \left\|(x-y) - \lambda A(x-y)\right\|^2 \\
&= \left\|x-y\right\|^2 - 2\lambda\langle x-y, A(x-y)\rangle + \lambda^2\left\|A(x-y)\right\|^2 \\
&\leqslant \left\|x-y\right\|^2\left(1 - 2\lambda\alpha + \lambda^2\left\|A\right\|^2\right) \\
&= c_\lambda^2\left\|x-y\right\|^2 .
\end{aligned}$$

□

Define the mapping T_λ by

$$T_\lambda x := T\left(\lambda b + (I - \lambda A)x\right) = T\left(\lambda b + G_\lambda x\right) \ , \ x \in H \ .$$

Since T is nonexpansive, it is easy to see that T_λ is a contraction with contraction coefficient c_λ . Therefore T_λ has a unique fixed point, denoted ξ_λ . Namely, we have

$$\xi_\lambda = T_\lambda \xi_\lambda = T\left(\lambda b + (I - \lambda A)\xi_\lambda\right) = T\left(\lambda b + G_\lambda \xi_\lambda\right) \ .$$

Our first result is the following theorem.

Theorem 3.1.

$$s\text{–}\lim_{\lambda \to 0} \xi_\lambda = x^* \ .$$

Proof. (1)

$$(\xi_\lambda) \text{ is bounded .}$$

We have

$$\begin{aligned}
\|\xi_\lambda - x^*\| &= \|T_\lambda \xi_\lambda - T x^*\| \\
&= \|T\left(\lambda b + (I - \lambda A)\xi_\lambda\right) - T x^*\| \\
&\leqslant \|(\lambda b + (I - \lambda A)\xi_\lambda) - x^*\| \\
&= \|(I - \lambda A)(\xi_\lambda - x^*) + \lambda(b - A x^*)\| \\
&\leqslant \|(I - \lambda A)(\xi_\lambda - x^*)\| + \lambda \|b - A x^*\| \\
&\leqslant \|I - \lambda A\| \, \|\xi_\lambda - x^*\| + \lambda \|b - A x^*\| \\
&\leqslant (1 - \alpha\lambda) \|\xi_\lambda - x^*\| + \lambda \|b - A x^*\| \ .
\end{aligned}$$

It turns out that

$$\|\xi_\lambda - x^*\| \leqslant \frac{1}{\alpha} \|b - A x^*\| \ .$$

Hence (ξ_λ) is bounded, let $M > 0$ satisfy

$$\|\xi_\lambda\| \leqslant M \text{ for all } \lambda > 0 \ .$$

(2)

$$\|T\xi_\lambda - \xi_\lambda\| \to 0 \text{ as } \lambda \to 0 \ .$$

This is because

$$\begin{aligned}\|T\xi_\lambda - \xi_\lambda\| &= \|T\xi_\lambda - T(\lambda b + (I - \lambda A)\xi_\lambda)\| \\ &\leqslant \|\xi_\lambda - (\lambda b + (I - \lambda A)\xi_\lambda)\| \\ &= \lambda \|b - A\xi_\lambda\| \\ &\leqslant \lambda (\|b\| + M \|A\|) \to 0 \text{ as } \lambda \to 0 \ .\end{aligned}$$

Consequently, by the demiclosedness of $I - T$, we conclude that if $\lambda_n \to 0$ and if $\xi_{\lambda_n} \rightharpoonup \xi$, then $(I - T)\xi = 0$, that is, $\xi \in Fix(T)$.

(3)

$$(\xi_\lambda) \text{ converges in norm to } x^*.$$

Toward this we assume that $\lambda_n \to 0$ and $\xi_n := \xi_{\lambda_n} \rightharpoonup \xi$. We shall prove that $\xi_n \to x^*$ by establishing the following inequality

$$\alpha \|\xi_n - x^*\|^2 \leqslant \langle (I - \lambda_n A)(\xi_n - x^*), b - Ax^* \rangle + \langle b - Ax^*, y_n - x^* \rangle \ , \tag{3.1}$$

where

$$y_n = \lambda_n b + (I - \lambda_n A)\xi_n \ .$$

Note that we can write

$$\xi_n = Ty_n \quad \text{and} \quad y_n - x_n = \lambda_n(b - Ax_n) \ .$$

Next we estimate $\|\xi_n - x^*\|^2$ as follows

$$\begin{aligned}\|\xi_n - x^*\|^2 &= \|Ty_n - Tx^*\|^2 \leqslant \|y_n - x^*\|^2 \\ &= \langle y_n - x^*, y_n - x^* \rangle \\ &= \langle (I - \lambda_n A)(\xi_n - x^*), y_n - x^* \rangle + \lambda_n \langle b - Ax^*, y_n - x^* \rangle \ .\end{aligned}$$

Since $y_n - x^* = (I - \lambda_n A)(\xi_n - x^*) + \lambda_n(b - Ax^*)$, we get

$$\begin{aligned}\|\xi_n - x^*\|^2 &\leqslant \|(I - \lambda_n A)(\xi_n - x^*)\|^2 + \langle (I - \lambda_n A)(\xi_n - x^*), \lambda_n(b - Ax^*) \rangle + \lambda_n \langle b - Ax^*, y_n - x^* \rangle \\ &\leqslant (1 - \alpha\lambda_n) \|\xi_n - x^*\|^2 + \lambda_n \langle (I - \lambda_n A)(\xi_n - x^*), b - Ax^* \rangle + \lambda_n \langle b - Ax^*, y_n - x^* \rangle \ .\end{aligned}$$

It follows that

$$\alpha \|\xi_n - x^*\|^2 \leqslant \langle (I - \lambda_n A)(\xi_n - x^*), b - Ax^* \rangle + \langle b - Ax^*, y_n - x^* \rangle \ ,$$

this is (3.1) .

By (2), we know $\xi \in Fix(T)$. As $\|y_n - \xi_n\| \to 0$, we also have $y_n \rightharpoonup \xi$.

Observe also as $\lambda_n \to 0$,

$$\begin{aligned}\|(I - \lambda_n A)(\xi_n - x^*) - (\xi_n - x^*)\| &\leqslant \lambda_n \|A(\xi_n - x^*)\| \\ &\leqslant \lambda_n(\|A\|(M + \|x^*\|)) \to 0 .\end{aligned}$$

We obtain $(I - \lambda_n A)(\xi_n - x^*) \rightharpoonup \xi - x^*$.

Therefore, upon taking the limit in (3.1) as $n \to \infty$, we get

$$\limsup_{n\to\infty} \|\xi_n - x^*\|^2 \leqslant \frac{2}{\alpha}\langle b - Ax^*, \xi - x^*\rangle . \tag{3.2}$$

The optimality condition for $\mathrm{Min}_C(\theta)$ is

$$\langle Ax^* - b, x - x^*\rangle \geqslant 0 \text{ for all } x \in C = Fix(T) .$$

Now, since $\xi \in Fix(T)$, (3.2) implies $\xi_n \to x^*$ in norm.

This also concludes the proof that $\xi_\lambda \to x^*$ in norm as $\lambda \to 0$. □

We next study an iterative method for finding the solution x^* . Our method generates a sequence $(x_n)_{n=0}^{\infty}$ iteratively as follows :
The initial guess $x_0 \in H$ is selected arbitrarily, and x_{n+1} is defined by the recursion process

$$x_{n+1} = T_{\lambda_n} x_n = T\left(\lambda_n b + (I - \lambda_n A)x_n\right) \ , \ \text{n} \geqslant 0 . \tag{3.3}$$

Theorem 3.2. *Assume* $0 < \lambda < \frac{2\alpha}{\|A\|^2}$ *and* $\{\lambda_n\}$ *satisfies the conditions*

(*i*) $\lim_{n\to\infty} \lambda_n = 0$,

(*ii*) $\sum_{n=0}^{\infty} \lambda_n = \infty$,

(*iii*) *either* $\sum_{n=0}^{\infty} |\lambda_{n+1} - \lambda_n| < \infty$ *or* $\lim_{n\to\infty} \frac{\lambda_n}{\lambda_{n+1}} = 1$.

Then (x_n) *converges strongly to* x^* .

Proof. (1)

$$(x_n) \text{ is bounded .}$$

Indeed, we have

$$\begin{aligned}\|x_{n+1} - x^*\| &= \|T_{\lambda_n} x_n - T x^*\| \\ &\leqslant \|T_{\lambda_n} x_n - T_{\lambda_n} x^*\| + \|T_{\lambda_n} x^* - T x^*\| \\ &\leqslant c_{\lambda_n} \|x_n - x^*\| + \lambda_n \|b - Ax^*\| \ ,\end{aligned}$$

where $c_{\lambda_n} \leqslant 1 - \frac{1}{2}\lambda_n \left(2\alpha - \lambda_n \|A\|^2\right)$.

As $\lambda_n \to 0$, we may assume $\lambda_n \|A\|^2 < \alpha$ for all n ; hence, $c_{\lambda_n} \leqslant 1 - \frac{\alpha}{2}\lambda_n$.

It follows that

$$\begin{aligned}\|x_{n+1} - x^*\| &\leqslant \left(1 - \frac{1}{2}\alpha\lambda_n\right) \|x_n - x^*\| + \frac{1}{2}\alpha\lambda_n \left(\frac{2\|b - Ax^*\|}{\alpha}\right) \\ &\leqslant \max\{\|x_n - x^*\| , \frac{2}{\alpha}\|b - Ax^*\|\} \ , \ \text{n} \geqslant 0 \ .\end{aligned}$$

By induction, we find

$$\|x_n - x^*\| \leqslant \max\{\|x_0 - x^*\| , \frac{2}{\alpha}\|b - Ax^*\|\} \ .$$

Hence, (x_n) is bounded. Let $M > 0$ satisfy

$$\begin{array}{rcl}\|x_n\| & \leqslant & M \ , \\ \|Ax_n\| & \leqslant & M \ , \qquad \text{for all } n \geqslant 0 \ . \\ \|b - Ax_n\| & \leqslant & M \ ,\end{array}$$

(2)

$$\|x_{n+1} - Tx_n\| \to 0 \ .$$

Indeed,

$$\begin{aligned}\|x_{n+1} - Tx_n\| &= \|T_{\lambda_n} x_n - T x_n\| \\ &= \|T(\lambda_n b + (I - \lambda_n A)x_n) - Tx_n\| \\ &\leqslant \|\lambda_n b + (I - \lambda_n A)x_n - x_n\| \\ &= \lambda_n \|b - Ax_n\| \leqslant M\lambda_n \to 0 \ .\end{aligned}$$

(3)

$$\|x_{n+1} - x_n\| \to 0 \ .$$

As a matter of fact, we deduce

$$\begin{aligned}
\|x_{n+1} - x_n\| &= \left\|T_{\lambda_n} x_n - T_{\lambda_{n-1}} x_{n-1}\right\| \\
&\leqslant \|T_{\lambda_n} x_n - T_{\lambda_n} x_{n-1}\| + \left\|T_{\lambda_n} x_{n-1} - T_{\lambda_{n-1}} x_{n-1}\right\| \\
&\leqslant \left(1 - \frac{\alpha}{2}\lambda_n\right) \|x_n - x_{n-1}\| \\
&\quad + \|T\left(\lambda_n b + (I - \lambda_n A)x_{n-1}\right) - T\left(\lambda_{n-1} b + (I - \lambda_{n-1} A)x_{n-1}\right) x_{n-1}\| \\
&\leqslant \left(1 - \frac{\alpha}{2}\lambda_n\right) \|x_n - x_{n-1}\| + |\lambda_n - \lambda_{n-1}| \, \|b - Ax_{n-1}\| \ .
\end{aligned}$$

Apply the lemma below to get (3) .

Lemma 3.3. *Let $(a_n)_{n \geqslant 0}$ be a sequence of positive real numbers such that*

$$a_{n+1} \leqslant (1 - \gamma_n) \, a_n + \gamma_n b_n \ , \quad n \geqslant 0 \ ,$$

where (b_n) is a sequence of real numbers, such that

$$(i) \ \sum_{n=0}^{\infty} \gamma_n = \infty \ ,$$

$$(ii) \ \text{either} \ \sum_{n=0}^{\infty} \gamma_n \, |b_n| < \infty \ \text{or} \ \limsup_{n \to \infty} b_n \leqslant 0 \ ,$$

then $\lim_{n \to \infty} a_n = 0$.

(4)

$$\limsup_{n \to \infty} \langle b - Ax^*, x_n - x^* \rangle \leqslant 0 \ .$$

To see this, take a subsequence $\left(x_{n_j}\right)$ of (x_n) , such that

$$\lim_{n \to \infty} \langle b - Ax^*, x_n - x^* \rangle = \lim_{j \to \infty} \langle b - Ax^*, x_{n_j} - x^* \rangle \ , \ \text{and} \ x_{n_j} \rightharpoonup \hat{x} \ .$$

This, together with Lemma 3.3, implies that (as $\hat{x} \in Fix(T)$)

$$\limsup_{n \to \infty} \langle b - Ax^*, x_n - x^* \rangle = \limsup_{n \to \infty} \langle b - Ax^*, \hat{x} - x^* \rangle \leqslant 0 \ .$$

(5)

$$x_n \to x^* \text{ in norm },$$

Set $y_n = \lambda_n b + (I - \lambda_n A)\, x_n$. Then $\|y_n - x_n\| \to 0$, so we also have by step (4)

$$\limsup_{n\to\infty} \langle b - Ax^*, y_n - x^* \rangle \leqslant 0 \ .$$

We obtain

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &= \|T_{\lambda_n} x_n - x^*\|^2 \\
&= \|T y_n - T x^*\|^2 \\
&\leqslant \|y_n - x^*\|^2 \\
&= \|\lambda_n b + (I - \lambda_n A) x_n - x^*\|^2 \\
&= \|(I - \lambda_n A)(x_n - x^*) + \lambda_n (b - Ax^*)\|^2 \\
&\leqslant \|(I - \lambda_n A)(x_n - x^*)\|^2 + 2\lambda_n \langle b - Ax^*, (I - \lambda_n A)(x_n - x^*) + \lambda_n (b - Ax^*) \rangle \\
&\leqslant (1 - \alpha\lambda_n)^2 \|x_n - x^*\|^2 + 2\lambda_n \langle b - Ax^*, y_n - x^* \rangle \\
&\leqslant (1 - \alpha\lambda_n) \|x_n - x^*\|^2 + \alpha\lambda_n \left(\frac{2}{\alpha} \langle b - Ax^*, y_n - x^* \rangle \right) \ .
\end{aligned}$$

Applying Lemma 3.3 to the case where

$$a_n := \|x_n - x^*\|^2 \ , \ \gamma_n := \alpha\lambda_n \ , \ b_n := \frac{2}{\alpha} \langle b - Ax^*, y_n - x^* \rangle \ ,$$

we conclude that $\|x_n - x^*\|^2 \to 0$. That is, $x_n \to x^*$ strongly . □

References

[1] Juan Peypouquet, Coupling the Gradient Method with a General Exterior Penalization Scheme for Convex Minimization, *J Optim Theory Appl.* 153 (2012), 123–138.

[2] John G. O'Hara, Paranjothi Pillay and Hong-Kun Xu, Iterative Approaches to Convex Minimization Problems, *Numerical Functional Analysis and Optimization.* 25(5&6) (2004), 531–546.

[3] O'Hara, J. G., Pillay, P. and Hong-Kun Xu, Iterative approaches to finding nearest common fixed points of nonexpansive mappings in Hilbert spaces, *Nonlinear Anal.* 54 (2003), 1417–1426.

[4] Hong-Kun Xu, An iterative approach to quadratic optimization, *J. Optimiz. Theory Appl.* 116 (2003), 659–678.

[5] Halpern, B., Fixed points of nonexpanding maps, *Bull. Amer. Math. Soc.* 73 (1967), 957–961.

[6] Wittmann, R., Approximation of fixed points of nonexpansive mappings, *Arch. Math.* 58 (1992), 486–491.

[7] Bauschke, H., The appoximation of fixed points of compositions of nonexpansive mappings in Hilbert space, *J. Math. Anal. Appl.* 202 (1996), 150–159.

[8] Deutsch, F., Yamada, I., Minimizing certain convex functions over the intersection of the fixed point sets of nonexpansive mappings, *Numer. Funct. Anal. Optim.* 19(1&2) (1998), 33–56.

[9] Yamada, I., Ogura, N., Yamashita, Y., Sakaniwa, K., Quadratic approximation of fixed points of nonexpansive mappings in Hilbert spaces, *Numer. Funct. Anal. Optimiz.* 19(1) (1998), 165–190.