CONVERGENCE THEOREMS FOR SOLVING MINIMIZATION PROBLEM

KUNRADA KANKAM

A Thesis Submitted to University of Phayao
in Partial Fulfillment of the Requirements
for the Master of Science Degree in Mathematics
December 2019

Thesis

Title

Convergence Theorems for Solving Minimization Problem

Submitted by Kunrada Kankam

Approved in partial fulfillment of the requirements for the

Master of Science Degree in Mathematics

University of Phayao

....Chairman

(Professor Dr.Suthep Suantai)

....Committee

(Associate Professor Dr.Prasit Cholamjiak)

....Committee

(Associate Professor Dr.Tanakit Thianwan)

....Committee

(Assistant Professor Dr.Damrongsak Yambangwai)

....Committee

(Assistant Professor Dr.Watcharaporn Cholamjiak)

Approved by

....

(Associate Professor Dr.Chayan Boonyarak)

Dean of School of Science

December 2019

ACKNOWLEDGEMENT

First of all, I would like to express my sincere appreciation to my supervisor, Associate Professor Dr.Prasit Cholamjiak for his primary idea, guidance and motivation which enable me to carry out my study successfully.

I gladly thank to the supreme committees, Professor Dr.Suthep Suantai, Associate Professor Dr.Tanakit Thianwan, Assistant Professor Dr.Damrongsak Yambangwai and Assistant Professor Dr.Watcharaporn Cholamjiak, for recommendation about my presentation, report and future works.

I also thank to all of my teachers for their previous valuable lectures that give me more knowledge during my study at Department of Mathematics, School of Science, University of Phayao.

I am thankful for all my friends with their help and warm friendship. Finally, my graduation would not be achieved without best wish from my parents, who help me for everything and always gives me greatest love, willpower and financial support until this study completion.

Kunrada Kankam

เรื่อง: ทฤษฎีบทการลู่เข้าสำหรับการแก้ปัญหาค่าต่ำสุด

ผู้วิจัย: กุลดา ขันคำ, วิทยานิพนธ์: วท.ม. (คณิตศาสตร์), มหาวิทยาลัยพะเยา, 2562

ประธานที่ปรึกษา: รองศาสตราจารย์ ดร.ประสิทธิ์ ช่อลำเจียก

กรรมการที่ปรึกษา: รองศาสตราจารย์ ดร.ธนกฤต เทียนหวาน, ผู้ช่วยศาสตราจารย์ ดร.ดำรงศักดิ์ แย้มบางหวาย, ผู้ช่วยศาสตราจารย์ ดร.วัชรภรณ์ ช่อลำเจียก

คำสำคัญ: วิธีการข้างหน้า–ข้างหลัง, การลู่เข้าแบบเข้ม, การหาค่าต่ำสุดเชิงคอนเวกซ์, ปริภูมิฮิลเบิร์ต

บทคัดย่อ

ปัญหาจริงมากมายทางด้านวิทยาศาสตร์ประยุกต์ วิศวกรรมศาสตร์ และ เศรษฐศาสตร์ สามารถแปลงให้อยู่ในรูปแบบของปัญหาค่าต่ำสุดเชิงคอนเวกซ์ของผลรวมของสองฟังก์ชันวัตถุประสงค์ เพื่อที่จะแก้ปัญหานี้วิธีการแยกข้างหน้า–ข้างหลังได้ถูกนำมาใช้สำหรับการวิเคราะห์การลู่เข้า อย่างไรก็ตามโดยทั่วไปเงื่อนไขความต่อเนื่องลิพชิทซ์ของเกรเดียนต์ของฟังก์ชันมักจะถูกกำหนดขึ้น ซึ่งเป็นสิ่งที่ยากในการคำนวณ นอกจากนี้ข้อสมมติฐานนี้ยังทำให้เกิดการลู่เข้าที่ช้าของอัลกอริทึม วัตถุประสงค์หลักของวิทยานิพนธ์นี้คือการปรับปรุงและพัฒนาวิธีการแยกแบบใหม่สำหรับการแก้ปัญหาค่าต่ำสุดเชิงคอนเวกซ์ อันดับแรกจะทำการพิสูจน์ทฤษฎีบทการลู่เข้าแบบเข้มของลำดับที่ก่อกำเนิดโดยวิธีการข้างหน้า–ข้างหลังโดยระเบียบวิธีการภาพฉายลูกผสมและวิธีการฉายภาพหดตัวภายใต้ปริภูมิฮิลเบิร์ต อันดับต่อมาจะทำการพิสูจน์ทฤษฎีบทการลู่เข้าแบบเข้มของลำดับที่ก่อกำเนิดโดยวิธีการข้างหน้า–ข้างหลังโดยระเบียบวิธีการประมาณแบบยืดหยุ่นภายใต้ปริภูมิฮิลเบิร์ต ในงานวิจัยนี้จะศึกษาขนาดขั้นแบบใหม่ของไลน์เสิร์ชสองรูปแบบที่แตกต่างกัน ข้อได้เปรียบหลักของอัลกอริทึมที่ได้พัฒนาขึ้นมาคือค่าคงที่ลิพชิทซ์ของเกรเดียนต์ของฟังก์ชันไม่จำเป็นต้องใช้ในการคำนวณ อันดับสุดท้ายการทดลองเชิงตัวเลขแสดงให้เห็นถึงประสิทธิภาพของวิธีการที่ได้นำเสนอในรูปแบบของการกู้คืนสัญญาณ ผลลัพธ์เชิงตัวเลขแสดงให้เห็นว่าวิธีการที่ได้ถูกนำเสนอมีอัตราการลู่เข้าที่ดีกว่าวิธีการอื่นที่เกี่ยวข้อง

Title: CONVERGENCE THEOREMS FOR SOLVING MINIMIZATION PROBLEM

Author: Kunrada Kankam, Thesis: M.Sc. (Mathematics), University of Phayao, 2019

Advisor: Associate Professor Dr.Prasit Cholamjiak

Committee: Associate Professor Dr.Tanakit Thianwan, Assistant Professor Dr.Damrongsak Yambangwai, Assistant Professor Dr.Watcharaporn Cholamjiak

Keywords: forward–backward method, strong convergence, convex minimization, Hilbert space.

ABSTRACT

Many real world problems in applied sciences, engineering and economics can be reformulated as the convex minimization problem of the sum of two objective functions. In order to solve this problem, the forward–backward splitting algorithm has been used for the convergence analysis. However, in general, the Lipschitz continuity condition on the gradient of functions is usually assumed which is not an easy task in computation. Moreover, this assumption leads to the slow convergence of algorithms. The main objective of this thesis is to improve and develop new splitting algorithms for solving convex minimization problem. First, strong convergence theorems of the sequences generated by the forward–backward algorithms using hybrid projection method and shrinking projection method are proved in Hilbert spaces. Second, strong convergence theorems of the sequence generated by the forward–backward algorithm using viscosity approximation method are proved in Hilbert spaces. The stepsizes studied in this thesis are defined by two different kinds of linesearches. The main advantage of our algorithms is that the Lipschitz constants of the gradient of functions do not require in computation. Finally, numerical experiments are given to show the efficiency of the proposed methods in signal recovery. Numerical results show that the proposed algorithms have a better convergence rate than other related algorithms.

LIST OF CONTENTS

LIST OF TABLES

Tables **Page**

LIST OF FIGURES

Figures **Page**

CHAPTER I

INTRODUCTION

Many real world problems in applied sciences, engineering and economics can be reformulated as the optimization problem (OP). An OP refers to the general problem of minimizing (or maximizing) objective function that are typically not differentiable at their minimizers. More generally, optimization includes finding best available values of some objective function given a defined domain, including a variety of different types of objective functions and different types of domains.

To solve minimization problems, researchers may use algorithm that terminate in finite number of steps, or iterative methods that converge to a solution, or heuristics that may provide approximate solutions to problem, one of the most important techniques in handling ill - posed problems and inverse problems. The Tikhonov regularization and proximal point methods are widely used to deal with one maximal monotone operator. The proximal point algorithm (PPA) initiated by Martinet in 1970 and subsequently studied by Rockafellar in 1976 is often referred. However, since the PPA does not necessarily converges strongly, many researchers have conducted worthwhile work on modifying the PPA so that the strong convergence is guaranteed, for examples, the relaxed proximal point algorithm (RPPA) and the contraction proximal point algorithm (CPPA). The splitting methods play a central role in the analysis and the numerical solution of such problems. The Forward-Backward and Douglas-Rachford splitting algorithms are classical methods for computing those reliable solutions. Due to its applications, there have been several modifications and generalizations of these methods suggested and invented independently for solving the problem in many different contexts. This tool plays an important role in the analysis and the numerical solution of convex optimization problems. The main concept of the proximal mapping technique is obtained by splitting in that the functions are used individually so as to yield an easily implementable algorithm, for example, the proximal point algorithm (Martinet, Rockafellar) is used to find a minimizer of a convex function, the forward-backward algorithm is used to find a minimizer of the sum of two convex functions and so on. One of the main advantages of these algorithms is that they can be used, without computation on the projection which is not an easy task in general, to minimize nondifferentiable objectives, such as those commonly encountered in sparse approximation and compressed sensing, or in hard-constrained problems as well as involving high-dimensional data. There have recently been many researchers extensively studied and developed this technique based on the proximity operators such as Wang (2000), Nakajo and Takahashi (2003), Combettes and Wajs (2005). However, many proximal point method usually assumed that the gradient is Lipschitz continuous and the step size is bounded below and less than some constants related to the Lipschitz constant, which is some how not known in practice. For this reason, it is our purpose to study and develop new algorithm for solving minimization problems.

Over the past few years, Bello Cruz and Nghia (2016) studied and developed proximal mapping technique for solving minimization problems by the proximal gradient algorithm using new linesearch technique for solving the convex minimization problem in Hilbert spaces. The main advantage of the proposed method is that the Lipschitz condition on the gradient of functions is dropped in computing. As reviewed, it is therefore the main objective in this research to develop and modify the numerical algorithms by using the proximal mapping technique and the linesearch rules for solving minimization problems and to establish some convergence theorems which admit less stringent and/or more constructive requirements on solving minimization problems. The main results established in this research can improve and generalize the corresponding results in this area and, of course, can be applied to solve major problems existed in science, engineering, economics and other related branches. Finally, to give some applications of minimization problem including its numerical experiments. The main results can improve and extend the corresponding results in this area and, can be applied to solve major problems existed in science.

CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

In this work, we study in solving the convex minimization problem which is modeled as the following form:

$$\min_{x \in H} f(x) + g(x), \tag{2.1.1}$$

where H is a real Hilbert space with the inner product $\langle \cdot, \cdot \rangle$, the induced norm $\| \cdot \|$ and $f, g : H \to \mathbb{R} \cup \{+\infty\}$ are two proper, lower-semicontinuous and convex functions in which f is Fréchet differentiable on an open set containing the domain of g. The solution set of problem (2.1.1) will be denoted by S_*. It is known that (2.1.1) relates to the following fixed point equation:

$$x = \operatorname{prox}_{\beta g}(x - \beta \nabla f(x)) \tag{2.1.2}$$

where β is a positive real number and prox_g is the proximal operator of g. Using this fixed point equation, one can define the following classical forward-backward algorithm:

$$x^{k+1} = \underbrace{\operatorname{prox}_{\beta_k g}}_{\text{backward step}} \underbrace{(x^k - \beta_k \nabla f(x^k))}_{\text{forward step}}, \tag{2.1.3}$$

where β_k is a suitable stepsize. This method includes, in particular, the proximal point algorithm [12, 21, 24, 27] and the gradient method [11, 30, 32]. Due to its wide applications, there have been modifications of (2.1.3) invented in the literature (see [6, 7, 10, 16, 17, 26]).

In 2003, Nakajo and Takahashi [23] introduced the following hybrid projection method and prove its strong convergence for finding a fixed point of a nonexpansive mapping T. Let C be a nonempty closed convex subset of a real Hilbert spaces H. They investigated the sequence (x^k) generated by: $x^0 \in C$ and

$$\begin{cases} y^k = \alpha_k x^k + (1-\alpha_k)Tx^k, \\ C_k = \{z \in C : \|y^k - z\| \le \|x^k - z\|\}, \\ Q_k = \{z \in C : \langle z - x^k, x^0 - x^k \rangle \le 0\}, \\ x^{k+1} = P_{C_k \cap Q_k}(x^0), \end{cases} \tag{2.1.4}$$

for every $k \in \mathbb{N} \cup \{0\}$, where $(\alpha_k) \subset [0, a]$ for some $a \in [0, 1)$. They proved that (x^k) converges strongly to a fixed point of T. Furthermore, Takahashi et al. [28] proposed the shrinking projection method which is defined by: $x^0 \in C$, $C_1 = C$, $x^1 = P_{C_1}(x^0)$ and

$$\begin{cases} y^k = \alpha_k x^k + (1-\alpha_k)Tx^k, \\ C_{k+1} = \{z \in C_k : \|y^k - z\| \le \|x^k - z\|\}, \\ x^{k+1} = P_{C_{k+1}}(x^0), \end{cases} \tag{2.1.5}$$

where $0 \le \alpha_k < a < 1$ for all $k \in \mathbb{N}$. It was proved that the sequence (x^k) generated by (2.1.5) converges strongly to a fixed point of a nonexpansive mapping T.

In 2000, Moudafi [22] introduced the viscosity approximation method for fixed point problem of nonexpansive mappings. To this end they associate to the initial problem, namely

$$\text{fine } x \in C \text{ such that } x = T(x),$$

where T is strongly nonexpansive, the following approximate well - posed problem

$$\text{find } x^k \in C \text{ such that } x^k = \frac{1}{1+\varepsilon_k}T(x^k) + \frac{\varepsilon_k}{1+\varepsilon_k}h(x^k),$$

where $\{\varepsilon_k\}$ is a sequence of positive real numbers having to go to zero and $h : X \to C$ is a contraction. When suppose $\sum_{k=1}^{+\infty} \varepsilon_k = +\infty$ and $\lim_{k\to\infty} |\frac{1}{\varepsilon_k} - \frac{1}{\varepsilon_{k-1}}| = 0$. Then, for all x_0, the sequence $\{x^k\}$ converges strongly to a fixed point of T. It is well known that this method establishes strong convergence.

In 2012, Lin and Takahashi [18] introduced the following modification:

Algorithm 2.1.1 :

Initialization Step. *Take* $x^0 \in H$

Iterative Step. *Give* x^k *and set*

$$x^{k+1} = a_k h(x^k) + (1 - a_k)\text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)),$$

where $h : H \to H$ *is a* $\rho-$*contraction for some* $\rho \in [0, 1)$, *i.e.* $\|h(x) - h(y)\| \leq \rho\|x - y\|$ *for all* $x, y \in H$ *and* ∇f *is a* $\nu-$*inverse strongly monotone with the following conditions:*

$$\lim_{n\to\infty} a_k = 0, \sum_{k=1}^{\infty} a_k = \infty, \sum_{k=1}^{\infty} |a_k - a_{k+1}| < \infty;$$

$$\sum_{k=1}^{\infty} |\alpha_k - \alpha_{k+1}| < \infty, \ 0 < b \leq \alpha_k \leq 2\nu$$

Stop Criteria. *If* $x^{k+1} = x^k$, *then stop.*

Recently, Wang and Wang [31] proposed the following forward-backward splitting method:

Algorithm 2.1.2 *Let arbitrary initial guess* $x^1 \in H$, *and generates* x^{k+1} *accord-*

ing to the recursion process,

$$x^{k+1} = a_k F(x^k) + b_k x^k + c_k \mathrm{prox}_{\beta_k g}(x^k - \beta_k \nabla f(x^k)), \tag{2.1.6}$$

where $(a_k) \subset (0,1)$, $(b_k) \subset (-2,1)$, $(c_k) \subset (0,2)$ *and* $a_k + b_k + c_k = 1$, *and* $F : H \to H$ *is a contraction.*

Theorem 2.1.3 *Let* (β_k) *be a sequence in* $(0, \frac{2}{L})$. *Suppose that the following conditions are satisfied:*

(i) $\lim\limits_{k\to\infty} a_k = 0$, $\sum\limits_{k=1}^{\infty} a_k = \infty$;

(ii) $\lim\limits_{k\to\infty} \dfrac{a_k}{c_k} = 0$;

(iii) $\limsup\limits_{k\to\infty} c_k < \dfrac{\frac{4}{L}}{\frac{2}{L} + \limsup\limits_{k\to\infty} \beta_k}$;

(iv) $0 < \liminf\limits_{k\to\infty} \beta_k \leq \limsup\limits_{k\to\infty} \beta_k < \dfrac{2}{L}$.

Then the sequence (x^k) *generated by (2.1.6) converges strongly to a minimizer of* $f + g$.

The forward-backward method based on iteration (2.1.3) has been studied by many authors: see e.g. [5, 7, 8, 10, 16, 19, 20, 25, 29]. However, it should be noted that the stepsize β_k usually depends on the Lipschitz assumption on the gradient of a function f. This leads to the difficulty since the Lipschitz constants are often unknown in general.

Recently Cruz and Nghia [3] investigated the forward-backward method using linesearch that eliminates the undesired Lipschitz assumption on the gradient of f and proved the weak convergence of sequences generated by the proposed algorithm to optimal solution as follows:

Linesearch 2.1.4 *Given* $x,\ \sigma > 0,\ \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$.

Input. *Set* $\beta = \sigma$ *and* $J(x,\beta) := \text{prox}_{\beta g}(x - \beta\nabla f(x))$ *with* $x \in \text{dom}g$.

While $\beta\|\nabla f(J(x,\beta)) - \nabla f(x)\| > \delta\|J(x,\beta) - x\|$

do $\beta = \theta\beta$.

End While

Output. β.

It was proved that Linesearch 2.1.4 is well - defined, i.e., this linesearch stops after finitely many steps. So it can be considered the following algorithm:

Algorithm 2.1.5 :

Initialization Step. *Take* $x^0 \in \text{dom}g,\ \sigma > 0,\ \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$

Iterative Step. *Give* x^k *and set*

$$x^{k+1} = \text{prox}_{\beta_k g}(x^k - \beta_k \nabla f(x^k)),$$

with $\beta_k :=$ *Linesearch 2.1.4* $(x^k, \sigma, \theta, \delta)$.

Stop Criteria. *If* $x^{k+1} = x^k$, *then stop.*

It was shown that the sequence generated by Algorithm 2.1.5 converges weakly to minimizers of $f + g$. Moreover, if the gradient of f is globally Lipschitz continuous on dom$g$ with a constant $L > 0$, then $\alpha_k \geq \min\{\sigma, \delta\theta/L\}$ for all $k \in \mathbb{N}$. However, their algorithms have only weak convergence in real Hilbert spaces. As pointed out, for example, by Bauschke and Combettes [2], the weak convergence of an iterative scheme is an unsatisfactory property in an infinite dimensional setting. Moreover, it is our academic interests to analyze the strong convergence using the linesearch technique.

In this research, inspired by Cruz and Nghia [3], we introduce new algorithm, based on Algorithm 2.1.5, hybrid projection method (2.1.4), shrinking projection method (2.1.5) and Algorithm 2.1.5, using the viscosity approximation method. We then prove the strong convergence theorems of the proposed methods. We also suggest a new linesearch which is different from Linesearch 2.1.4. Combining this linesearch and the forward-backward method, we also prove its strong convergence in Hilbert spaces. Finally, some numerical experiments, in signal recovery, are provided to show the efficiency and the implementation of our algorithms. The report shows that our algorithms can be applied to solve the compressed sensing in the frequency domain. Moreover, it is discovered that the forward-backward algorithm using new linesearch has a better convergence than others in comparison. The main advantage is that our schemes do not require the information of the Lipschitz constant of the gradient of functions which makes the proposed algorithm more practical for computing.

CHAPTER III

PRELIMINARIES

3.1 Fundamentals

In this section, we provide some basic concepts, definitions and lemmas which will be used in the sequel.

Definition 3.1.1 (Metric space) Let X be a nonempty set and $d : X \times X \to [0, \infty)$ be a function. Then d is called a *metric* on X if the following properties hold:

1. $d(x, y) \geq 0$ for all $x, y \in X$;
2. $d(x, y) = 0$ if and only if $x = y$ for all $x, y \in X$;
3. $d(x, y) = d(y, x)$ for all $x, y \in X$;
4. $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in X$.

The value of metric d at (x, y), we write $d(x, y)$, is called *distance* between x and y, and the ordered pair (X, d) is called a *metric space*.

Example 3.1.2 In real line $\mathbb{R}$, define

$$d(x, y) = |x - y| \tag{3.1.1}$$

for all $x, y \in \mathbb{R}$. Then $(\mathbb{R}, d)$ is a metric space.

Example 3.1.3 In euclidean plane $\mathbb{R}^2$, define

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2} \tag{3.1.2}$$

where $x = (\xi_1, \xi_2), y = (\eta_1, \eta_2) \in \mathbb{R}^2$. Then $(\mathbb{R}^2, d)$ is a metric space.

Example 3.1.4 In euclidean space $\mathbb{R}^k$, define

$$d(x,y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^2 + ... + (\xi_k - \eta_k)^2} \tag{3.1.3}$$

where $x = (\xi_1, \xi_2, \xi_3, ..., \xi_k), y = (\eta_1, \eta_2, \eta_3, ..., \eta_k) \in \mathbb{R}^k$. Then $(\mathbb{R}^k, d)$ is a metric space.

Example 3.1.5 Let X be the set of all bounded sequences of complex numbers; that is every element of X is a complex sequence

$$x = (\xi_1, \xi_2, ...)$$

such that $|\xi_j| \leq c_x$ for all $j = 1, 2, ...$ and c_x is a real number which may depend on x, but does not depend on j and define

$$d(x,y) = \sup_{j \in \mathbb{N}} |\xi_j - \eta_j| \tag{3.1.4}$$

where $y = (\eta_j) \in X$ and $\mathbb{N} = 1, 2, ...$ Then (X, d) is a metric space.

Definition 3.1.6 (Open and Closed sets) Let (X, d) be a metric space. A subset $U \subseteq X$ is open if for every $x \in X$ there exists $r > 0$ such that $B(x, r) \subseteq U$. A set U is closed if its complement, $X \setminus U$, is open.

Definition 3.1.7 (Convergent sequence) A sequence (x^k) in a metric space X is said to be convergent to $x \in \mathbb{R}$ if for each $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x^k, x) < \varepsilon$ for all $k > N$. In this case, we write $x^k \to x$.

Definition 3.1.8 (Cauchy sequence) A sequence (x^k) in a metric space X is said to be *Cauchy* if for each $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x^j, x^k) < \epsilon$ for all $j, k > N$.

Theorem 3.1.9 *Let M be a nonempty subset of a metric space X. Then M is closed if and only if there exists a sequence $\{x^k\} \subseteq M$ and $x^k \to x$ implies that*

$x \in M$.

Definition 3.1.10 (Bounded sequence) A sequence (x^k) in X is bounded if there exists $M > 0$ such that $\|x^k\| \leq M$ for all $k \in \mathbb{N}$.

Definition 3.1.11 (Nonexpansive mapping) Let (X, d) be a metric space. Then a map $T : X \to X$ is said to be *nonexpansive* if

$$d(T(x), T(y)) \leq d(x, y)$$

for all $x, y \in X$.

Definition 3.1.12 (Contractive mapping) Let (X, d) be a metric space. Then a map $T : X \to X$ is said to be *contractive* if there exists $k \in [0, 1)$ such that

$$d(T(x), T(y)) \leq kd(x, y)$$

for all $x, y \in X$.

Definition 3.1.13 (Fixed point) Let X be a nonempty set and $T : X \to X$. We say that $x \in X$ is a fixed point of T if

$$T(x) = x \tag{3.1.5}$$

and denote by $Fix(T)$ the set of all fixed points of T.

Theorem 3.1.14 (The Banach contraction principle) *Let X be a complete metric space and let T be a contraction of X into itself. Then T has a unique fixed point.*

Definition 3.1.15 (Vector space) A vector space or linear space X over the field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) is a set X together with an internal binary operation $(+)$ called addition and a scalar multiplication carrying (α, x) in $\mathbb{K} \times X$ to αx in X satisfying the following statements for all $x, y, z \in X$ and $\alpha, \beta \in \mathbb{K}$:

1. $x + y = y + x$;
2. $(x + y) + z = x + (y + z)$;
3. there exists an element $0 \in X$ call the *zero vector* of X such that $x + 0 = x$ for all $x \in X$;
4. for every element $x \in X$, there exists an element $-x \in X$ called *the additive inverse* or *the negative* of x such that $x + (-x) = 0$;
5. $\alpha(x + y) = \alpha x + \alpha y$;
6. $(\alpha + \beta)x = \alpha x + \beta y$;
7. $(\alpha\beta)x = \alpha(\beta x)$;
8. $1 \cdot x = x$.

The elements of a vector space X are called vectors, and the elements of $\mathbb{K}$ are called scalars.

Example 3.1.16 In euclidean space $\mathbb{R}^k$, define

$$\begin{aligned} x + y &= (\xi_1 + \eta_1, \xi_2 + \eta_2, \xi_3 + \eta_3, ..., \xi_k + \eta_k) \\ \alpha x &= (\alpha\xi_1, \alpha\xi_2, \alpha\xi_3, ..., \alpha\xi_k) \end{aligned}$$

where $x = (\xi_1, \xi_2, \xi_3, ..., \xi_k), y = (\eta_1, \eta_2, \eta_3, ..., \eta_k) \in \mathbb{R}^k$ and $\alpha \in \mathbb{R}$. Then, space $\mathbb{R}^k$ is a real vector space.

Definition 3.1.17 (Convex set) Let C be a subset of a linear space X. Then C is said to be *convex* if $(1 - \lambda)x + \lambda y \in C$ for all $x, y \in C$ and all scalar $\lambda \in [0, 1]$.

Example 3.1.18 1. Every subspace of vector space is convex.

2. $\overline{B}(x; r) = \{x : \|x\| \leq r\}$ is convex.
3. $[0, 1]^K = [1, 0] \times [1, 0] \times ... \times [1, 0]$ is convex in $\mathbb{R}^k$.

Proposition 3.1.19 *Let C be a subset of a linear space X. Then C is convex if and only if $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_k x_k \in C$ for any finite set $\{x_1, x_2, ..., x_k\} \subseteq C$ and scalars $\lambda_i \geq 0$ with $\lambda_1 + \lambda_2 + ... + \lambda_k = 1$.*

Definition 3.1.20 (Convex function) Let X be a linear space and $f : X \to (-\infty, \infty]$ be a function. Then f is said to be *convex* if $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in X$ and $\lambda \in [0, 1]$.

Definition 3.1.21 (Proper function) Let function $f : X \to (-\infty, \infty]$. Then f is said to be *proper* if there exists $x \in X$ with $f(x) < \infty$.

Example 3.1.22 1. $f(x) = |x|^p$ where $p \geq 1$ is a convex function in $\mathbb{R}$.

2. $f(x) = x^3 - x^2$ is a convex function in $[\frac{1}{3}, \infty)$.

3. $f(x) = x \log x$ is a convex function in $\mathbb{R}^+$.

Definition 3.1.23 (Normed space) Let X be a norm linear space over field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) and $\|\cdot\| : X \to \mathbb{R}^+$ be a function. Then $\|\cdot\|$ is said to be a norm if the following properties hold:

1. $\|x\| \geq 0$, and $\|x\| = 0 \Leftrightarrow x = 0$;
2. $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and $\alpha \in \mathbb{K}$;
3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

The ordered pair $(X, \|\cdot\|)$ is called a *normed space*.

Example 3.1.24 $\mathbb{R}^k$ is a normed space with the following norms:

$$\begin{aligned}
\|x\|_1 &= \sum_{i=1}^{k} |x_i| \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k; \\
\|x\|_p &= \Big(\sum_{i=1}^{k} |x_i|^p\Big)^{1/p} \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k \text{ and } p \in (1, \infty); \\
\|x\|_\infty &= \max_{1 \leq i \leq k} |x_i| \text{ for all } x = (x_1, x_2, .., x_k) \in \mathbb{R}^k.
\end{aligned}$$

Example 3.1.25 Let $X = l_1$, the linear space whose elements consist of all absolutely convergent sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_1 = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \sum_{i=1}^{\infty} |x_i| < \infty\}.$$

Then l_1 is a normed space with the norm defined by $\|x\|_1 = \sum_{i=1}^{\infty} |x_i|$.

Example 3.1.26 Let $X = l_p$ $(1 < p < \infty)$ be the linear space whose elements consist of all p-summable sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_p = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \sum_{i=1}^{\infty} |x_i|^p < \infty\}.$$

Then l_p is a normed space with the norm defined by $\|x\|_p = (\sum_{i=1}^{\infty} |x_i|^p)^{1/p}$.

Example 3.1.27 Let $X = l_\infty$, the linear space whose elements consist of all bounded sequences $(x_1, x_2, ..., x_i, ...)$ of scalars ($\mathbb{R}$ *or* $\mathbb{C}$),

$$l_\infty = \{x : x = (x_1, x_2, ..., x_i, ...) \text{ and } \{x_i\}_{i=1}^{\infty} \text{ is bounded}\}.$$

Then l_∞ is a normed space with the norm defined by $\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$.

Definition 3.1.28 (Completeness) The space X is said to be *complete* if every Cauchy sequence in X converges.

Example 3.1.29 The Euclidean space $\mathbb{R}^k$ is complete with

$$d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2 + (\xi_3 - \eta_3)^2 + ... + (\xi_k - \eta_k)^2} \tag{3.1.6}$$

where $x = (\xi_1, \xi_2, \xi_3, ..., \xi_k), y = (\eta_1, \eta_2, \eta_3, ..., \eta_k) \in \mathbb{R}^k$.

Example 3.1.30 The sequence space l_∞ is complete.

Example 3.1.31 The sequence space l_p is complete.

Definition 3.1.32 (Inner product space) An inner product space is a vector space X with an inner product defined on X. Here, an inner product on X is a mapping of $X \times X$ into the scalar field $\mathbb{K}$ of X; that is, with every pair of vectors x and y there is associated a scalar which is written by $\langle x, y\rangle$ and called the *inner product* of x and y, such that for all vectors x, y, z and scalars α we have

(IP1) $\langle x, x\rangle \geq 0$;

(IP2) $\langle x, x\rangle = 0 \Leftrightarrow x = 0$;

(IP3) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$;

(IP4) $\langle x, y\rangle = \overline{\langle y, x\rangle}$;

(IP5) $\langle x + y, z\rangle = \langle x, z\rangle + \langle z, y\rangle$.

Example 3.1.33 The function $\langle\cdot,\cdot\rangle : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ defined by

$$\langle x, y\rangle = \sum_{i=1}^{k} x_i y_i \text{ for all } x = (x_1, x_2, ..., x_k),\ y = (y_1, y_2, ..., y_k) \in \mathbb{R}^k \tag{3.1.7}$$

is an inner product on $\mathbb{R}^k$. In this case $\mathbb{R}^k$ with this inner product is called real Euclidean k-space.

Example 3.1.34 Let $\mathbb{C}^k$ be the set of k-tuples of complex numbers. Then the function $\langle\cdot,\cdot\rangle : \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ defined by

$$\langle x, y\rangle = \sum_{i=1}^{k} x_i \overline{y_i} \text{ for all } x = (x_1, x_2, ..., x_k),\ y = (y_1, y_2, ..., y_k) \in \mathbb{C}^k \tag{3.1.8}$$

is an inner product on $\mathbb{C}^k$. In this case $\mathbb{C}^k$ with this inner product is called complex Euclidean k-space.

Example 3.1.35 Let l_2 be the set of all sequences of complex numbers $(a_1, a_2, \ldots, a_i, \ldots)$ with $\sum_{i=1}^{\infty} |a_i|^2 < \infty$. Then the function $\langle\cdot,\cdot\rangle : l_2 \times l_2 \to \mathbb{C}$ defined by

$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i \overline{y_i} \text{ for all } x = \{x_i\}_{i=1}^{\infty}, y = \{y_i\}_{i=1}^{\infty} \in l_2 \tag{3.1.9}$$

is an inner product on l_2.

Definition 3.1.36 (Hilbert space) An inner product space which is complete with respect to the induced norm is called a *Hilbert space*.

Example 3.1.37 The Euclidean space $\mathbb{R}^k$ is a Hilbert space with inner product defined by

$$\langle x, y\rangle = x^1y^1 + x^2y^2 + ... + x^ky^k$$

where $x = (x^1, x^2, ..., x^k),\ y = (y^1, y^2, ..., y^k) \in \mathbb{R}^k$.

Example 3.1.38 The space l_2 is a Hilbert space with inner product defined by

$$\langle x, y\rangle = \sum_{j=1}^{\infty} x_j\overline{y_j},$$

where $x, y \in l_2$.

Proposition 3.1.39 (The Cauchy-Schwarz inequality) *Let X be an inner product space. Then the following holds:*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle \text{ for all } x, y \in X, \tag{3.1.10}$$

i.e.,

$$|\langle x, y\rangle| \leq \|x\|\|y\| \text{ for all } x, y \in X. \tag{3.1.11}$$

Definition 3.1.40 (Bounded linear operator) Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. The operator T is said to be *bounded* if there is a real number $c > 0$ such that for all $x \in X$,

$$\|Tx\| \leq c\|x\|. \tag{3.1.12}$$

Definition 3.1.41 (Level set of convex function) Let $f : H \to \mathbb{R}$ be a convex function with the domain H. Then, for any $\lambda \in \mathbb{R}$, the set

$$V_\lambda = \{x \in H | f(x) \leq \lambda\} \tag{3.1.13}$$

Definition 3.1.42 A sequence (x^k) in a Hilbert space H is said to converge weakly to a point x in H if

$$\langle x^k, y\rangle \to \langle x, y\rangle \tag{3.1.14}$$

for all $y \in H$ and denote that $x^k \rightharpoonup x$.

Definition 3.1.43 (Contraction mapping) Let H be a real Hilbert space and C be a nonempty subset of H. Then a map $F : C \to C$ is said to be *contraction* if there exists $k \in [0, 1)$ such that

$$\|F(x) - F(y)\| \leq k\|x - y\|,$$

for all $x, y \in C$.

Definition 3.1.44 (Nonexpansive mapping) Let H be a real Hilbert space and C be a nonempty subset of H. A mapping $T : C \to C$ is said to be *nonexpansive* if

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

A mapping $T : C \to C$ is said to be *firmly nonexpansive* if, for all $x, y \in C$,

$$\langle x - y, Tx - Ty\rangle \geq \|Tx - Ty\|^2. \tag{3.1.15}$$

The operator $I - P_C$ is also *firmly nonexpansive*, where I denotes the identity operator, *i.e.*, for any $x, y \in H$,

$$\langle (I - P_C)x - (I - P_C)y, x - y\rangle \geq \|(I - P_C)x - (I - P_C)y\|^2. \tag{3.1.16}$$

In a real Hilbert space, we know that for any point $x \in H$, there exists a unique point $P_C x \in C$ such that

$$\|x - P_C x\| \leq \|x - y\|, \forall y \in C.$$

Here 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

$$\langle x-y, P_Cx-P_Cy\rangle \geq \|P_Cx-P_Cy\|^2, \tag{3.1.17}$$

for all $x, y \in H$. Furthermore, P_Cx is characterized by the property

$$\langle x-P_Cx, P_Cx-y\rangle \geq 0, \tag{3.1.18}$$

for all $y \in C$. Moreover, we know that

$$\|y-P_Cx\|^2+\|x-P_Cx\|^2 \leq \|x-y\|^2,\ \forall x, y \in H. \tag{3.1.19}$$

We also know that all Hilbert space has the Kadec-Klee property, that is, (x^k) converges weakly to x and $\|x^k\| \to \|x\|$ imply x^k converges strongly to x.

Lemma 3.1.45 *[13] Assume (s^k) is a sequence of nonnegative real numbers such that*

$$s^{k+1} \leq (1-\omega_k)s^k+\omega_k\chi_k, k \geq 1 \tag{3.1.20}$$

and

$$s^{k+1} \leq s^k-\psi_k+\varphi_k, \tag{3.1.21}$$

where (ω_k) is a sequence in $(0,1), (\psi_k)$ is a sequence of nonnegative real numbers and $(\chi_k), (\varphi_k)$ are real sequences such that

(1) $\sum_{k=1}^{\infty}\omega_k = \infty$,

(2) $\lim_{k\to\infty}\varphi_k = 0$,

(3) $\lim_{n\to\infty}\psi_{k_n} = 0$ *implies* $\limsup_{n\to\infty}\chi_{k_n} \leq 0$ *for any subsequence of real numbers (k_n) of (k).*

Then $\lim_{k\to\infty} s^k = 0$.

Definition 3.1.46 Let H be a real Hilbert space and let $f : H \to R$, function f

is said to be lower semi-continuous at x if $x^k \to x$, then

$$f(x) \leq \liminf_{k\to\infty} f(x^k).$$

Definition 3.1.47 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function, The proximal operator $\text{prox}_f : \mathbb{R}^n \to \mathbb{R}^n$ of f is defined by

$$\text{prox}_f(v) = \arg\min_x (f(x) + (1/2)\|x - v\|_2^2),$$

and the proximal operator of the scalar function αf, where $\alpha > 0$, which can be expressed as

$$\text{prox}_{\alpha f}(v) = \arg\min_x (f(x) + (1/2\alpha)\|x - v\|_2^2),$$

then $\text{prox}_{\alpha f}$ is call the proximal operator of f with parameter α.

Definition 3.1.48 Let H be a real Hilbert space and let $h : H \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous (l.s.c.), and convex function. The subdifferential of h at x is defined by

$$\partial h(x) = \{v \in H : \langle v, y - x\rangle + h(x) \leq h(y), y \in H\}.$$

Example 3.1.49 The real line R, $f : R \to R$ by $f(x) = |x|$. The subdifferential,

$$\partial f(x) = \begin{cases} -1 & \text{if } x > 0, \\ [-1, 1] & \text{if } x = 0, \\ 1 & \text{if } x < 0. \end{cases} \tag{3.1.22}$$

Proof. For $x \in R$, we have that

$$z \in \partial f(x) \Leftrightarrow |y| - |x| \geq z(y - x)\ \forall y \in R. \tag{3.1.23}$$

We consider the three cases of $x > 0$, $x < 0$ and $x = 0$. Let $x > 0$. If $y > x$, then we have from (3.1.23) that $y - x \geq z(y - x)$ and hence $1 \geq z$. For y with $0 < y < x$, we have from (3.1.23) that $y - x \geq z(y - x)$ and hence $1 \leq z$. So, we have $z = 1$. Let $x < 0$. As in the proof of $x > 0$, we have $z = -1$. In the case of $x = 0$, we have from (3.1.23) that $|y| \geq zy$. If $y > 0$, then we have $y \geq zy$ and hence $1 \geq z$. If $y < 0$, then we have $-y \geq zy$ and hence $-1 \leq z$. So, we have $-1 \leq z \leq 1$. Then, we have (3.1.22) □

Recall that an element $g \in H$ is said to be a subgradient of $f : H \to \mathbb{R}$ at x if

$$f(z) \geq f(x) + \langle g, z - x \rangle, \ \forall z \in H.$$

Fact 3.1.50 *[[2], Proposition 17.2] Let $h : H \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower-semicontinuous and convex function. Then, for $x \in \mathrm{dom}h$ and $y \in H$, the following hold:*

(1) $f'(x; y)$ exists and

$$f'(x; y) = \inf_{\alpha \in R^{+} \cup (\infty)} \frac{f(x + \alpha y) - f(x)}{\alpha}.$$

(2) $h'(x; y - x) + h(x) \leq h(y)$.

Lemma 3.1.51 *[4] The subdifferential operator ∂h is maximal monotone. Moreover, the graph of ∂h, $\mathrm{Gph}(\partial h) = \{(x, v) \in H \times H : v \in \partial h(x)\}$ is demiclosed, i.e., if the sequence $(x^k, v^k) \subset \mathrm{Gph}(\partial h)$ satisfies that $(x^k)_{k \in \mathbb{N}}$ converges weakly to x and $(v^k)_{k \in \mathbb{N}}$ converges strongly to v, then $(x, v) \in \mathrm{Gph}(\partial h)$.*

Let us recall the proximal operator $\mathrm{prox}_g : H \to \mathrm{dom}g$ with $\mathrm{prox}_g(z) = (\mathrm{I} + \partial g)^{-1}(z), z \in H$. Here I denotes the identity operator. It is well - known that the proximal operator is single - valued with full domain. It is also known that

$$\frac{z - \mathrm{prox}_{\alpha g}(z)}{\alpha} \in \partial g(\mathrm{prox}_{\alpha g}(z)) \text{ for all } z \in H, \ \alpha > 0. \tag{3.1.24}$$

CHAPTER IV

MAIN RESULTS

4.1 Hybrid forward-backward algorithms using linesearch rule for minimization problem

In this section, we propose the forward-backward splitting algorithm using the projection algorithm and prove the strong convergence theorem.

Following [3], we assume that two below conditions hold:

(A1) $f, g : H \to \mathbb{R} \cup \{+\infty\}$ are two proper, lower-semicontinuous and convex functions with $\mathrm{dom} g \subseteq \mathrm{dom} f$ and $\mathrm{dom} g$ is nonempty, closed and convex.

(A2) The function f is Fréchet differentiable on an open set containing $\mathrm{dom} g$. The gradient ∇f is uniformly continuous on any bounded subset of $\mathrm{dom} g$ and maps any bounded subset of $\mathrm{dom} g$ to a bounded set in H.

Algorithm 4.1.1 *(step 0) Choose* $x^0 \in \mathrm{dom} g$*, take* $\delta \in (0, \frac{1}{2})$*,* $\sigma > 0$ *and* $\theta \in (0, 1)$.

(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\begin{aligned} &\alpha_k \|\nabla f(\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))) - \nabla f(x^k)\| \\ \leq\ &\delta \|\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)) - x^k\|. \end{aligned} \tag{4.1.1}$$

(step 2) Set

$$y^k = \mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{4.1.2}$$

(step 3) Compute

$$C_k = \{x_* \in \mathrm{dom} g : \|y^k - x_*\| \leq \|x^k - x_*\|\}$$

and

$$Q_k = \{x_* \in \text{dom}g : \langle x_* - x^k, x^0 - x^k\rangle \leq 0\}. \tag{4.1.3}$$

(step 4) Compute

$$x^{k+1} = P_{C_k \cap Q_k}(x^0). \tag{4.1.4}$$

(step 5) Set $k \leftarrow k+1$, and go to (step 1).

Denote S_* by the solution set of (2.1.1) and assume that S_* is nonempty.

Theorem 4.1.2 *Let H be a real Hilbert space. Assume that there exists $\alpha > 0$ such that $\alpha_k \geq \alpha > 0$. Then the sequence $(x^k)_{k=0}^{\infty}$ generated by Algorithm 4.1.1 converges strongly to $\bar{x} = P_{S_*}(x^0)$.*

Proof. We divide our proof into four steps.

Step 1 Show that $(x_k)_{k=0}^{\infty}$ is well - defined and $S_* \subset C_k \cap Q_k, \forall k \geq 0$. For each $x \in \text{dom}g$, we see that

$$\begin{aligned} \|y^k - x\| \leq \|x^k - x\| \quad &\leftrightarrow \quad \|y^k\|^2 - 2\langle x, y^k\rangle \leq \|x^k\|^2 - 2\langle x, x^k\rangle \\ &\leftrightarrow \quad 2\langle x, x^k - y^k\rangle \leq \|x^k\|^2 - \|y^k\|^2 \\ &\leftrightarrow \quad \langle x, x^k - y^k\rangle \leq \frac{1}{2}[\|x^k\|^2 - \|y^k\|^2]. \end{aligned} \tag{4.1.5}$$

Therefore C_k is closed and convex for all $k \geq 0$. Moreover, it is easy to see that Q_k is closed and convex for all $k \geq 0$. Therefore, $C_k \cap Q_k$ is closed and convex for all $k \geq 0$. Using (3.1.24) and (4.1.2), we observe that

$$\frac{x^k - y^k}{\alpha_k} - \nabla f(x^k) = \frac{x^k - \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))}{\alpha_k} - \nabla f(x^k) \in \partial g(y^k).$$

The convexity of g gives

$$g(x) - g(y^k) \geq \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle, \forall x \in \mathrm{dom} g. \tag{4.1.6}$$

The convexity of f also implies

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle, \forall x \in \mathrm{dom} f,\ \ y \in \mathrm{dom} g. \tag{4.1.7}$$

From (4.1.6) and (4.1.7) with any $x \in \mathrm{dom} g \subseteq \mathrm{dom} f$ and $y = x^k \in \mathrm{dom} g$, we have

$$\begin{aligned}
(f+g)(x) &\geq f(x^k) + g(y^k) + \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle + \langle \nabla f(x^k), x - x^k \rangle \\
&= f(x^k) + g(y^k) + \langle \nabla f(y^k), y^k - x^k \rangle \\
&\quad + \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle + \langle \nabla f(x^k) - \nabla f(y^k), y^k - x^k \rangle \\
&\geq f(x^k) + g(y^k) + \langle \nabla f(y^k), y^k - x^k \rangle \\
&\quad + \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle - \|\nabla f(x^k) - \nabla f(y^k)\| \|y^k - x^k\| \\
&\geq f(x^k) + g(y^k) + \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle - \frac{\delta}{\alpha_k} \|x^k - y^k\|^2 \\
&\quad + \langle \nabla f(y^k), y^k - x^k \rangle,
\end{aligned}$$

where the last inequality follows from the linesearch (4.1.1). Hence we obtain

$$\begin{aligned}
&\langle x^k - y^k, y^k - x \rangle \\
\geq\ & \alpha_k [f(x^k) + g(y^k) - (f+g)(x) + \langle \nabla f(y^k), y^k - x^k \rangle] \\
& - \delta \|x^k - y^k\|^2.
\end{aligned}$$

Replacing $x = x^k$ and $y = y^k$ in (4.1.7), we have $f(x^k) - f(y^k) \geq \langle \nabla f(y^k), x^k - y^k \rangle$.

So, we get

$$\langle x^k - y^k, y^k - x\rangle \;\; \geq \;\; \alpha_k[(f+g)(y^k) - (f+g)(x)] - \delta\|x^k - y^k\|^2.$$

Since $2\langle x^k - y^k, y^k - x\rangle = \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - x\|^2$, by (4.1.8), it follows that

$$\begin{aligned} & \|y^k - x\|^2 \\ \leq \;\; & \|x^k - x\|^2 - 2\alpha_k[(f+g)(y^k) - (f+g)(x)] - (1-2\delta)\|x^k - y^k\|^2 \end{aligned} \tag{4.1.8}$$

Let $x_* \in S_*$ and set $x = x_*$ in (4.1.8). Hence we have

$$\|y^k - x_*\| \leq \|x^k - x_*\|. \tag{4.1.9}$$

Thus $x_* \in C_k, \forall k \geq 0$. Therefore, $S_* \subset C_k, \forall k \geq 0$. For $k = 0$, we have that $x^0 \in \text{dom}g$ and $Q_0 = \text{dom}g$ and hence $S_* \subset C_0 \cap Q_0$. Assume that x^n is given and $S_* \subset C_n \cap Q_n$ for some $n \in \{0, 1, 2, ...\}$. Since S_* is nonempty, $C_n \cap Q_n$ is nonempty, closed and convex. So there exists a unique element $x^{n+1} \in C_n \cap Q_n$ such that $x^{n+1} = P_{C_n \cap Q_n}(x^0)$. This gives

$$\langle x_* - x^{n+1}, x^0 - x^{n+1}\rangle \leq 0, \;\forall x_* \in C_n \cap Q_n. \tag{4.1.10}$$

Since $S_* \subset C_n \cap Q_n$, in particular, we obtain

$$\langle x_* - x^{n+1}, x^0 - x^{n+1}\rangle \leq 0, \;\forall x_* \in S_*. \tag{4.1.11}$$

This implies that $S_* \subset Q_{n+1}$. By induction we conclude that, $S_* \subset C_k \cap Q_k, \forall k \geq 0$ and thus $(x^k)_{k=0}^{\infty}$ is well - defined.

Step 2 Show that $(x^k)_{k=0}^{\infty}$ is bounded. From (4.1.3), we see that

$$\langle x_* - x^k, x^0 - x^k \rangle \leq 0, \forall x_* \in Q_k.$$

This implies that $x^k = P_{Q_k}(x^0)$. Then we have

$$\|x^k - x^0\| \leq \|x^0 - x_*\|, \forall x_* \in Q_k.$$

Since $S_* \subset Q_k$, it follows that

$$\|x^k - x^0\| \leq \|x^0 - x_*\|, \forall x_* \in S_*. \tag{4.1.12}$$

In particular, since $x^{k+1} \in Q_k$,

$$\|x^k - x^0\| \leq \|x^{k+1} - x^0\|. \tag{4.1.13}$$

By (4.1.12) and (4.1.13), we obtain $\lim_{k\to\infty} \|x^k - x^0\|$ exists. Hence $(x^k)_{k=0}^{\infty}$ is bounded.

Step 3 Show that $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$. By (3.1.19) and the fact that $x^k = P_{Q_k}(x^0)$, we see that

$$\|x^{k+1} - x^k\|^2 \leq \|x^{k+1} - x^0\|^2 - \|x^k - x^0\|^2.$$

Since $\lim_{k\to\infty} \|x^k - x^0\|$ exists, it follows that $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.

Step 4 Show that $\lim_{k\to\infty} x^k = \bar{x}$, where $\bar{x} = P_{S_*}(x^0)$. From (4.1.3), $x^{k+1} \in C_k$ and Step 3, we see that

$$\|y^k - x^{k+1}\| \leq \|x^k - x^{k+1}\| \to 0, \; k \to \infty.$$

Hence we obtain

$$\begin{aligned}\|y^k - x^k\| &\leq \|y^k - x^{k+1}\| + \|x^{k+1} - x^k\| \\ &\to 0,\ k \to \infty. \qquad (4.1.14)\end{aligned}$$

Since $(x^k)_{k=0}^{\infty}$ is bounded, the set of its weak accumulation point is nonempty. Take any weak accumulation point ω of (x^k). So there is a subsequence $(x^{k_n})_{n=0}^{\infty}$ of $(x^k)_{k=0}^{\infty}$ weakly converging to ω. We get from (4.1.14) and assumption (A2) that

$$\lim_{n\to\infty} \|\nabla f(y^{k_n}) - \nabla f(x^{k_n})\| = 0. \qquad (4.1.15)$$

Since $y^{k_n} = \text{prox}_{\alpha_{k_n} g}(x^{k_n} - \alpha_{k_n} \nabla f(x^{k_n}))$, it follows from (3.1.24) that

$$\frac{x^{k_n} - \alpha_{k_n} \nabla f(x^{k_n}) - y^{k_n}}{\alpha_{k_n}} \in \partial g(y^{k_n})$$

which implies that

$$\frac{x^{k_n} - y^{k_n}}{\alpha_{k_n}} + \nabla f(y^{k_n}) - \nabla f(x^{k_n}) \in \nabla f(y^{k_n}) + \partial g(y^{k_n}) \subseteq \partial (f+g)(y^{k_n}). \quad (4.1.16)$$

From (4.1.14), (4.1.15) and (4.1.16), we conclude that $\omega \in S_*$ by Lemma 3.1.51. If $\bar{x} = P_{S_*}(x^0)$, it then follows from (4.1.12), the fact that $\omega \in S_*$ and the lower semicontinuity of the norm that,

$$\begin{aligned}\|x^0 - \bar{x}\| &\leq \|x^0 - \omega\| \\ &\leq \liminf_{n\to\infty} \|x^0 - x^{k_n}\| \\ &\leq \limsup_{n\to\infty} \|x^0 - x^{k_n}\| \\ &\leq \|x^0 - \bar{x}\|. \qquad (4.1.17)\end{aligned}$$

Hence we obtain $\lim\limits_{n\to\infty} \|x^{k_n} - x^0\| = \|x^0 - \omega\| = \|x^0 - \bar{x}\|$. This yields $x^{k_n} \to \omega =$

$\bar{x}, n \to \infty$. It follows that (x^k) converges weakly to $\bar{x}$. So we have

$$\begin{aligned} \|x^0 - \bar{x}\| &\leq \liminf_{n\to\infty} \|x^0 - x^k\| \\ &\leq \limsup_{n\to\infty} \|x^0 - x^k\| \\ &\leq \|x^0 - \bar{x}\|. \end{aligned} \tag{4.1.18}$$

This shows that $\lim_{n\to\infty} \|x^k - x^0\| = \|x^0 - \bar{x}\|$. From $x^k \rightharpoonup \bar{x}$, we also have $x^k - x^0 \rightharpoonup \bar{x} - x^0$. Since H satisfies the Kadec-Klee property, it follows that $x^k - x^0 \to \bar{x} - x^0$. Therefore $x^k \to \bar{x}$ as $k \to \infty$. This completes the proof. □

Next, we introduce another version of the forward-backward algorithm based on the shrinking projection method.

Algorithm 4.1.3 *(step 0) Set* $C_0 = \mathrm{dom}g$, *choose* $x^0 \in \mathrm{dom}g$, *take* $\delta \in (0, \frac{1}{2})$, $\sigma > 0$ *and* $\theta \in (0, 1)$.

(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k \|\nabla f(y^k) - \nabla f(x^k)\| \leq \delta \|y^k - x^k\|. \tag{4.1.19}$$

(step 2) Set

$$y^k = \mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{4.1.20}$$

(step 3) Compute

$$C_{k+1} = \{x_* \in C_k : \|y^k - x_*\| \leq \|x^k - x_*\|\}. \tag{4.1.21}$$

(step 4) Compute

$$x^{k+1} = P_{C_{k+1}}(x^0). \tag{4.1.22}$$

(step 5) Set $k \leftarrow k + 1$, *and go to (step 1).*

Theorem 4.1.4 *Let* H *be a real Hilbert space. Assume that there exists* $\alpha > 0$

such that $\alpha_k \geq \alpha > 0$. *Then the sequence* $(x^k)_{k=0}^{\infty}$ *generated by Algorithm 4.1.3 converges strongly to* $\bar{x} = P_{S_*}(x^0)$.

Proof. We divide our proof into five steps.

Step 1 Show that $P_{C_{k+1}}(x^0)$ is well - defined and $S_* \subseteq C_{k+1}, \forall k \geq 0$. Similar to Step 1 in Theorem 4.1.2, we can show that C_{k+1} is closed and convex, $\forall k \geq 0$. Also, we can show that

$$\|x^k - x^0\| \leq \|x^0 - x_*\|, \forall x_* \in C_k.$$

Thus, if $x_* \in S_*$, then we have $x_* \in C_{k+1}$. So $S_* \subseteq C_{k+1}$ and $P_{C_{k+1}}(x^0)$ is well - defined.

Step 2 Show that $\lim_{k\to\infty} \|x^k - x^0\|$ exists. From $x^k = P_{C_k}x^0$, $C_{k+1} \subset C_k$ and $x^{k+1} \in C_k$, $\forall k \geq 1$, we get

$$\|x^k - x^0\| \leq \|x^{k+1} - x^0\|, \forall k \geq 0.$$

On the other hand, since $S_* \subset C_k$, we obtain

$$\|x^k - x^0\| \leq \|x_* - x^0\|, \forall x_* \in S_*.$$

It follows that the sequence (x^k) is bounded and nondecreasing. Therefore, $\lim_{k\to\infty} \|x^k - x^0\|$ exists.

Step 3 Show that $x^k \to \bar{x}$ as $k \to \infty$. For $l > k$, by the definition of C_k, we see that $x^l = P_{C_l}(x^0) \in C_l \subset C_k$. So we obtain

$$\|x^l - x^k\|^2 \leq \|x^l - x^0\|^2 - \|x^k - x^0\|^2.$$

From Step 2, we have $(x^k)_{k=0}^{\infty}$ is a Cauchy sequence. Hence, $x^k \to \bar{x}$ as $k \to \infty$.

Step 4 Show that $\bar{x} \in S_*$. From Step 3, we see that

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0.$$

Since $x^{k+1} \in C_{k+1} \subset C_k$, we have

$$\|y^k - x^{k+1}\| \leq \|x^k - x^{k+1}\| \to 0,\ k \to \infty.$$

It follows that

$$\begin{aligned} \|y^k - x^k\| &\leq \|y^k - x^{k+1}\| + \|x^{k+1} - x^k\| \\ &\to 0,\ k \to \infty. \end{aligned} \tag{4.1.23}$$

We get from (4.1.23) and assumption (A2) that

$$\lim_{k\to\infty} \|\nabla f(y^k) - \nabla f(x^k)\| = 0. \tag{4.1.24}$$

Since $y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))$, it follows from (3.1.24) that

$$\frac{x^k - \alpha_k \nabla f(x^k) - y^k}{\alpha_k} \in \partial g(y^k)$$

which implies that

$$\frac{x^k - y^k}{\alpha_k} + \nabla f(y^k) - \nabla f(x^k) \in \nabla f(y^k) + \partial g(y^k) \subseteq \partial(f+g)(y^k). \tag{4.1.25}$$

From (4.1.23), (4.1.24) and (4.1.25), we have $\bar{x} \in S_*$. by Lemma 3.1.51

Step 5 Show that $\bar{x} = P_{S_*}(x^0)$. Since $x^k = P_{C_k}(x^0)$ and $S_* \subset C_k$, we obtain

$$\langle x^0 - x^k, x^k - x_* \rangle \geq 0, \forall x_* \in S_*. \tag{4.1.26}$$

By taking the limit in (4.1.26), we obtain

$$\langle x^0 - \bar{x}, \bar{x} - x_* \rangle \geq 0, \forall x_* \in S_*. \tag{4.1.27}$$

This shows that $\bar{x} = P_{S_*}(x^0)$. We thus complete the proof. □

4.2 Strong convergence of the forward-backward splitting algorithms via linesearches

Algorithm 4.2.1 *Let* $F : \text{domg} \to \text{domg}$ *be a contraction. Let* $\sigma > 0,\ \theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$, *take* $x^0 \in$ domg *and*

$$y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)) \tag{4.2.1}$$

where $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k \|\nabla f(y^k) - \nabla f(x^k)\| \leq \delta \|y^k - x^k\|. \tag{4.2.2}$$

Construct x^{k+1} *by*

$$x^{k+1} = a_k F(x^k) + (1 - a_k) y^k. \tag{4.2.3}$$

Lemma 4.2.2 *[3] The linesearch (4.2.2) stops after finitely many steps.*

Theorem 4.2.3 *Let* $(x^k)_{k\in\mathbb{N}}$ *and* (α_k) *be sequences generated by Algorithm 4.2.1. Suppose that there exists* $\alpha > 0$ *such that* $\alpha_k \geq \alpha$ *for all* $k \in \mathbb{N}$ *and* $(a_k) \subset (0,1)$ *such that*

$$\lim_{k\to\infty} a_k = 0 \text{ and } \sum_{k=1}^{\infty} a_k = \infty,$$

then the sequence $(x^k)_{k\in\mathbb{N}}$ *converges strongly to a point* $x_* = P_{S_*}F(x_*)$.

Proof. Using (3.1.24) and (4.2.1), we see that

$$\frac{x^k - y^k}{\alpha_k} - \nabla f(x^k) = \frac{x^k - \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))}{\alpha_k} - \nabla f(x^k) \in \partial g(y^k).$$

From the convexity of g, we have

$$g(x) - g(y^k) \geq \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle, \forall x \in \text{dom} g. \tag{4.2.4}$$

Also the convexity of f implies

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle, \forall x \in \text{dom} f,\ y \in \text{dom} g. \tag{4.2.5}$$

Combining (4.2.4) and (4.2.5) with any $x \in \text{dom} g \subseteq \text{dom} f$ and $y = x^k \in \text{dom} g$, we obtain

$$\begin{aligned}
& g(x) - g(y^k) + f(x) - f(x^k) \\
\geq\ & \langle \frac{x^k - y^k}{\alpha_k} - \nabla f(x^k), x - y^k \rangle + \langle \nabla f(x^k), x - x^k \rangle \\
=\ & \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle + \langle \nabla f(x^k) - \nabla f(y^k), y^k - x^k \rangle \\
& + \langle \nabla f(y^k), y^k - x^k \rangle \\
\geq\ & \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle - \|\nabla f(x^k) - \nabla f(y^k)\| \|y^k - x^k\| \\
& + \langle \nabla f(y^k), y^k - x^k \rangle \\
\geq\ & \frac{1}{\alpha_k} \langle x^k - y^k, x - y^k \rangle - \frac{\delta}{\alpha_k} \|x^k - y^k\|^2 + \langle \nabla f(y^k), y^k - x^k \rangle,
\end{aligned}$$

where the last inequality follows from (4.2.2). It then follows that

$$\begin{aligned}
& \langle x^k - y^k, y^k - x \rangle \\
\geq\ & \alpha_k [f(x^k) + g(y^k) - (f + g)(x) + \langle \nabla f(y^k), y^k - x^k \rangle] \\
& - \delta \|x^k - y^k\|^2.
\end{aligned}$$

Replacing $x = x^k$ and $y = y^k$ in (4.2.5), we have $f(x^k)-f(y^k) \geq \langle \nabla f(y^k), x^k - y^k \rangle$. We obtain

$$\langle x^k - y^k, y^k - x \rangle \geq \alpha_k[(f+g)(y^k) - (f+g)(x)] - \delta \|x^k - y^k\|^2. \quad (4.2.6)$$

Using $2\langle x^k - y^k, y^k - x\rangle = \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - x\|^2$, we get by (4.2.6) that

$$\|y^k - x\|^2 \leq \|x^k - x\|^2 - (1-2\delta)\|x^k - y^k\|^2 - 2\alpha_k[(f+g)(y^k) - (f+g)(x)]. \quad (4.2.7)$$

Let $x_* = P_S F(x_*)$. Then we have, by (4.2.7)

$$\|y^k - x_*\|^2 \leq \|x^k - x_*\|^2 - (1-2\delta)\|x^k - y^k\|^2. \quad (4.2.8)$$

Now, we will show that $(x^k)_{k\in\mathbb{N}}$ is bounded. Using (4.2.8), we get

$$\begin{aligned}
\|x^{k+1} - x_*\| &= \|a_k F(x^k) + (1-a_k)y^k - x_*\| \\
&\leq a_k\|F(x^k) - x_*\| + (1-a_k)\|y^k - x_*\| \\
&\leq a_k\|F(x^k) - x_*\| + (1-a_k)\|x^k - x_*\| \\
&\leq a_k\|F(x^k) - F(x_*)\| + a_k\|F(x_*) - x_*\| + (1-a_k)\|x^k - x_*\| \\
&\leq a_k c\|x^k - x_*\| + a_k\|F(x_*) - x_*\| + (1-a_k)\|x^k - x_*\| \\
&= (1 - a_k(1-c))\|x^k - x_*\| + a_k\|F(x_*) - x_*\|. \quad (4.2.9)
\end{aligned}$$

By induction, we can show that $(x^k)_{k\in\mathbb{N}}$ is bounded. On the other hand, we see that

$$\begin{aligned}
\|x^{k+1} - x_*\|^2 &= \langle x^{k+1} - x_*, x^{k+1} - x_* \rangle \\
&= \langle a_k(F(x^k) - x_*) + (1-a_k)(y^k - x_*), x^{k+1} - x_* \rangle \\
&\quad +(1-a_k)\langle y^k - x_*, x^{k+1} - x_* \rangle
\end{aligned}$$

$$
\begin{aligned}
&\leq a_k\|F(x^k)-F(x_*)\|\|x^{k+1}-x_*\|+a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
&\quad +(1-a_k)\|y^k-x_*\|\|x^{k+1}-x_*\| \\
&\leq \frac{a_k}{2}(\|F(x^k)-F(x_*)\|^2+\|x^{k+1}-x_*\|^2) \\
&\quad +a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
&\quad +\frac{(1-a_k)}{2}(\|y^k-x_*\|^2+\|x^{k+1}-x_*\|^2) \\
&\leq \frac{a_kc}{2}\|x^k-x_*\|^2+\frac{a_k}{2}\|x^{k+1}-x_*\|^2 \\
&\quad +a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
&\quad +\frac{(1-a_k)}{2}(\|y^k-x_*\|^2+\|x^{k+1}-x_*\|^2). \qquad (4.2.10)
\end{aligned}
$$

Using (4.2.8) and (4.2.10), we then have

$$
\begin{aligned}
\|x^{k+1}-x_*\|^2 &\leq \frac{a_kc}{2}\|x^k-x_*\|^2+\frac{a_k}{2}\|x^{k+1}-x_*\|^2+a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
&\quad +\frac{(1-a_k)}{2}(\|x^k-x_*\|^2-(1-2\delta)\|x^k-y^k\|^2+\|x^{k+1}-x_*\|^2).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\|x^{k+1}-x_*\|^2 &\leq (1-a_k(1-c))\|x^k-x_*\|^2-(1-2\delta)(1-a_k)\|x^k-y^k\|^2 \\
&\quad +2a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle. \qquad (4.2.11)
\end{aligned}
$$

In order to use Lemma 3.1.45, we set

$$
\begin{aligned}
s^k &= \|x^k-x_*\|^2 \\
\varphi_k &= 2a_k\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
\chi_k &= \frac{2}{(1-c)}\langle F(x_*)-x_*,x^{k+1}-x_*\rangle \\
\psi_k &= (1-2\delta)(1-a_k)\|x^k-y^k\|^2 \\
\omega_k &= a_k(1-c).
\end{aligned}
$$

So (4.2.11) reduces to the inequalities

$$s^{k+1} \leq (1-\omega_k)s^k + \omega_k \chi_k, k \geq 1 \tag{4.2.12}$$

$$s^{k+1} \leq s^k - \psi_k + \varphi_k. \tag{4.2.13}$$

Let (k_n) be a subsequence of (k) and suppose that $\lim_{n\to\infty} \psi_{k_n} = 0$. Then we have $\|x^{k_n} - y^{k_n}\| \to 0$ as $n \to \infty$. Also we obtain,

$$\begin{aligned} \|x^{k_n+1} - y^{k_n}\| &= \|a^{k_n} F(x^{k_n}) + (1-a^{k_n})y^{k_n} - y^{k_n}\| \\ &= a^{k_n}\|F(x^{k_n}) - y^{k_n}\| \\ &\to 0, \text{as } n \to \infty. \end{aligned} \tag{4.2.14}$$

Since $(x^{k_n})_{n\in\mathbb{N}}$ is bounded, the set of its weak accumulation points is nonempty. Take any weak accumulation point $\bar{x}$ of $(x^{k_n})_{n\in\mathbb{N}}$. So there is a subsequence $(x^{k_{n_i}})_{i\in\mathbb{N}}$ of $(x^{k_n})_{n\in\mathbb{N}}$ weakly converging to $\bar{x}$. We get from Assumption (A2) that

$$\lim_{i\to\infty} \|\nabla f(y^{k_{n_i}}) - \nabla f(x^{k_{n_i}})\| = 0. \tag{4.2.15}$$

Since $y^{k_{n_i}} = \text{prox}_{\alpha_{k_{n_i}} g}(x^{k_{n_i}} - \alpha_{k_{n_i}} \nabla f(x^{k_{n_i}}))$, it follows from (3.1.24) that

$$\frac{x^{k_{n_i}} - \alpha_{k_{n_i}} \nabla f(x^{k_{n_i}}) - y^{k_{n_i}}}{\alpha_{k_{n_i}}} \in \partial g(y^{k_{n_i}}) \tag{4.2.16}$$

which implies that

$$\begin{aligned} \frac{x^{k_{n_i}} - y^{k_{n_i}}}{\alpha_{k_{n_i}}} + \nabla f(y^{k_{n_i}}) - \nabla f(x^{k_{n_i}}) &\in \nabla f(y^{k_{n_i}}) + \partial g(y^{k_{n_i}}) \\ &\subseteq \partial(f+g)(y^{k_{n_i}}). \end{aligned} \tag{4.2.17}$$

Passing $i \to \infty$, by Lemma 3.1.51 and since $\|x^{k_{n_i}} - y^{k_{n_i}}\| \to 0$, we have $\bar{x} \in S_*$.

It follows that

$$\begin{aligned}\limsup_{n\to\infty}\langle F(x_*) - x_*, x^{k_n} - x_*\rangle &= \lim_{i\to\infty}\langle F(x_*) - x_*, x^{k_{n_i}} - x_*\rangle \\ &= \langle F(x_*) - x_*, \bar{x} - x_*\rangle \leq 0.\end{aligned}$$

We see that

$$\begin{aligned}\|x^{k_n+1} - x^{k_n}\| &\leq \|x^{k_n+1} - y^{k_n}\| + \|y^{k_n} - x^{k_n}\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{4.2.18}$$

From (4.2.18), we have

$$\limsup_{n\to\infty}\langle F(x_*) - x_*, x^{k_n+1} - x_*\rangle \leq 0. \tag{4.2.19}$$

Hence we get $\limsup_{n\to\infty} \chi_{k_n} \leq 0$. Using Lemma 3.1.45, we conclude that the sequence (x^k) converges strongly to $x_* = P_{S_*}F(x_*)$. □

We next introduce a new linesearch which is different from the previous linesearches.

Linesearch 4.2.4 *Given* $x \in \text{dom}g$, $\sigma > 0$, $\theta \in (0, 1)$ *and* $\delta \in (0, \frac{1}{2})$. *For* $i = 0, 1, 2, \ldots$, *set*

$$L(x, i) = \text{prox}_{\sigma\theta^i g}(x - \sigma\theta^i \nabla f(x))$$

and

$$S(x, i) = \text{prox}_{\sigma\theta^i g}(L(x, i) - \sigma\theta^i \nabla f(L(x, i))).$$

If

$$
\begin{aligned}
& 2\sigma\theta^i \max\{\|\nabla f(S(x,i)) - \nabla f(L(x,i))\|, \|\nabla f(L(x,i)) - \nabla f(x)\|\} \\
\leq \quad & \delta(\|S(x,i) - L(x,i)\| + \|L(x,i) - x\|) \qquad (4.2.20)
\end{aligned}
$$

then $\gamma = \sigma\theta^i$

Else $i = i + 1$.

We next show that Linesearch 4.2.4 is well - defined.

Lemma 4.2.5 *The Linesearch 4.2.4 stops after finitely many steps.*

Proof. If $x \in S_*$, then $x = \text{prox}_{\sigma g}(x - \sigma\nabla f(x)) = L(x, 0)$. It follows that $S(x, 0) = x$ and the linesearch stops with zero step and hence $\gamma = \sigma$. If $x \notin S_*$, then

$$
\begin{aligned}
& 2\sigma\theta^i \max\{\|\nabla f(S(x,i)) - \nabla f(L(x,i))\|, \|\nabla f(L(x,i)) - \nabla f(x)\|\} \\
& > \delta(\|S(x,i) - L(x,i)\| + \|L(x,i) - x\|). \qquad (4.2.21)
\end{aligned}
$$

So, we have as $i \to \infty$, $\|S(x,i) - L(x,i)\| \to 0$ and $\|L(x,i) - x\| \to 0$. By (A2), we see that $\|\nabla f(S(x,i)) - \nabla f(L(x,i))\| \to 0$ and $\|\nabla f(L(x,i)) - \nabla f(x)\| \to 0$ as $i \to \infty$. So, by (4.2.21), we have $\frac{\|x - L(x,i)\|}{\sigma\theta^i} \to 0$ as $i \to \infty$. Using (3.1.24), we have

$$
\frac{x - \sigma\theta^i\nabla f(x) - L(x,i)}{\sigma\theta^i} \in \partial g(L(x,i)).
$$

Hence $\frac{x - L(x,i)}{\sigma\theta^i} \in \partial g(L(x,i)) + \nabla f(x)$. So, as $i \to \infty$, we have $0 \in \partial g(x) + \nabla f(x)$ by Lemma 3.1.51. Thus $x \in S_*$ which is a contradiction. This completes the proof. □

Using this linesearch, we propose the following algorithm:

Algorithm 4.2.6 *Let* $F : \mathrm{dom}g \to \mathrm{dom}g$ *be a contraction. Let* $\sigma > 0$, $\theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$, *take* $x^0 \in$ dom$g$ *and*

$$y^k = \mathrm{prox}_{\gamma_k g}(x^k - \gamma_k \nabla f(x^k)), \tag{4.2.22}$$

where $\gamma_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\begin{aligned} & 2\gamma_k \max\{\|\nabla f(\mathrm{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k))) - \nabla f(y^k)\|, \|\nabla f(x^k) - \nabla f(y^k)\|\} \\ \leq\ & \delta(\|\mathrm{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k)) - y^k\| + \|x^k - y^k\|). \end{aligned} \tag{4.2.23}$$

Construct x^{k+1} *by*

$$x^{k+1} = a_k F(x^k) + (1 - a_k)\mathrm{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k)). \tag{4.2.24}$$

Theorem 4.2.7 *Let* $(x^k)_{k\in\mathbb{N}}$ *and* (γ_k) *be sequences generated by Algorithm 4.2.6. Suppose that there exists* $\gamma > 0$ *such that* $\gamma_k \geq \gamma$ *for all* $k \in \mathbb{N}$ *and* $(a_k) \subset (0,1)$ *such that*

$$\lim_{k\to\infty} a_k = 0 \text{ and } \sum_{k=1}^{\infty} a_k = \infty,$$

then the sequence $(x^k)_{k\in\mathbb{N}}$ *converges strongly to a point* $x_* = P_{S_*}F(x_*)$.

Proof. We set

$$z^k = \mathrm{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k)). \tag{4.2.25}$$

Using (3.1.24) and (4.2.22), we see that

$$\frac{x^k - y^k}{\gamma_k} - \nabla f(x^k) = \frac{x^k - \mathrm{prox}_{\gamma_k g}(x^k - \alpha_k \nabla f(x^k))}{\gamma_k} - \nabla f(x^k) \in \partial g(y^k).$$

By the convexity of g, it follows that

$$g(x) - g(y^k) \geq \langle \frac{x^k - y^k}{\gamma_k} - \nabla f(x^k), x - y^k \rangle, \forall x \in \mathrm{dom}g. \tag{4.2.26}$$

Also, from (3.1.24) and (4.2.25), we have

$$\frac{y^k - z^k}{\gamma_k} - \nabla f(y^k) = \frac{y^k - \text{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k))}{\gamma_k} - \nabla f(y^k) \in \partial g(z^k).$$

By the convexity of g, we also have

$$g(x) - g(z^k) \geq \langle \frac{y^k - z^k}{\gamma_k} - \nabla f(y^k), x - z^k \rangle, \forall x \in \text{dom}g. \quad (4.2.27)$$

We see that

$$f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle, \forall x \in \text{dom}f,\ y \in \text{dom}g. \quad (4.2.28)$$

For any $x \in \text{dom}g \subseteq \text{dom}f$ and $y = x^k$ in (4.2.28), we get

$$f(x) - f(x^k) \geq \langle \nabla f(x^k), x - x^k \rangle. \quad (4.2.29)$$

Also if $y = y^k$ in (4.2.28), then we get

$$f(x) - f(y^k) \geq \langle \nabla f(y^k), x - y^k \rangle. \quad (4.2.30)$$

So from (4.2.26), (4.2.27), (4.2.29) and (4.2.30), we have

$$\begin{aligned}
& g(x) - g(y^k) + g(x) - g(z^k) + f(x) - f(x^k) + f(x) - f(y^k) \\
\geq\ & \langle \frac{x^k - y^k}{\gamma_k} - \nabla f(x^k), x - y^k \rangle + \langle \frac{y^k - z^k}{\gamma_k} - \nabla f(y^k), x - z^k \rangle \\
& + \langle \nabla f(x^k), x - x^k \rangle + \langle \nabla f(y^k), x - y^k \rangle \\
=\ & \frac{1}{\gamma_k} \langle x^k - y^k, x - y^k \rangle + \langle \nabla f(x^k), y^k - x \rangle + \frac{1}{\gamma_k} \langle y^k - z^k, x - z^k \rangle \\
& + \langle \nabla f(y^k), z^k - x \rangle + \langle \nabla f(x^k), x - x^k \rangle + \langle \nabla f(y^k), x - y^k \rangle \\
=\ & \frac{1}{\gamma_k} [\langle x^k - y^k, x - y^k \rangle + \langle y^k - z^k, x - z^k \rangle] + \langle \nabla f(x^k), y^k - x^k \rangle \\
& + \langle f(y^k), z^k - y^k \rangle
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\gamma_k}[\langle x^k - y^k, x - y^k\rangle + \langle y^k - z^k, x - z^k\rangle] + \langle \nabla f(x^k) - \nabla f(y^k), y^k - x^k\rangle \\
&\quad +\langle \nabla f(y^k), y^k - x^k\rangle + \langle \nabla f(y^k) - \nabla f(z^k), z^k - y^k\rangle + \langle \nabla f(z^k), z^k - y^k\rangle \\
&\geq \frac{1}{\gamma_k}[\langle x^k - y^k, x - y^k\rangle + \langle y^k - z^k, x - z^k\rangle] - \|\nabla f(x^k) - \nabla f(y^k)\|\|y^k - x^k\| \\
&\quad +\langle \nabla f(y^k), y^k - x^k\rangle - \|\nabla f(y^k) - \nabla f(z^k)\|\|z^k - y^k\| \\
&\quad +\langle \nabla f(z^k), z^k - y^k\rangle. \qquad (4.2.31)
\end{aligned}
$$

Using (4.2.28), we obtain

$$
\begin{aligned}
&g(x) - g(y^k) + g(x) - g(z^k) + f(x) - f(x^k) + f(x) - f(y^k) \\
&\geq \frac{1}{\gamma_k}[\langle x^k - y^k, x - y^k\rangle + \langle y^k - z^k, x - z^k\rangle] - \|\nabla f(x^k) - \nabla f(y^k)\|\|y^k - x^k\| \\
&\quad +f(y^k) - f(x^k) - \|\nabla f(y^k) - \nabla f(z^k)\|\|z^k - y^k\| + f(z^k) - f(y^k). \qquad (4.2.32)
\end{aligned}
$$

So we have

$$
\begin{aligned}
&\frac{1}{\gamma_k}[\langle x^k - y^k, y^k - x\rangle + \langle y^k - z^k, z^k - x\rangle] \\
&\geq (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) - \|\nabla f(x^k) \\
&\quad -\nabla f(y^k)\|\|y^k - x^k\| - \|\nabla f(y^k) - \nabla f(z^k)\|\|z^k - y^k\|. \qquad (4.2.33)
\end{aligned}
$$

Using (4.2.23), we obtain

$$
\begin{aligned}
&\frac{1}{\gamma_k}[\langle x^k - y^k, y^k - x\rangle + \langle y^k - z^k, z^k - x\rangle] \\
&\geq (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) \\
&\quad -\frac{\delta}{2\gamma_k}(\|z^k - y^k\| + \|x^k - y^k\|)\|y^k - x^k\| \\
&\quad -\frac{\delta}{2\gamma_k}(\|z^k - y^k\| + \|x^k - y^k\|)\|z^k - y^k\| \\
&= (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) \\
&\quad -\frac{\delta}{2\gamma_k}(\|z^k - y^k\|\|y^k - x^k\| + \|y^k - x^k\|^2) \\
&\quad -\frac{\delta}{2\gamma_k}(\|z^k - y^k\|^2 + \|x^k - y^k\|\|z^k - y^k\|)
\end{aligned}
$$

$$
\begin{aligned}
&= (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) - \frac{\delta}{2\gamma_k}\|y^k - x^k\|^2 \\
&\quad -\frac{\delta}{2\gamma_k}\|z^k - y^k\|^2 - \frac{\delta}{\gamma_k}\|x^k - y^k\|\|z^k - y^k\| \\
&\geq (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) - \frac{\delta}{2\gamma_k}\|y^k - x^k\|^2 \\
&\quad -\frac{\delta}{2\gamma_k}\|z^k - y^k\|^2 - \frac{\delta}{2\gamma_k}\|x^k - y^k\|^2 - \frac{\delta}{2\gamma_k}\|z^k - y^k\|^2 \\
&= (f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x) - \frac{\delta}{\gamma_k}\|y^k - x^k\|^2 \\
&\quad -\frac{\delta}{\gamma_k}\|z^k - y^k\|^2. \qquad (4.2.34)
\end{aligned}
$$

We know that

$$2\langle x^k - y^k, y^k - x\rangle = \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - x\|^2$$

and

$$2\langle y^k - z^k, z^k - x\rangle = \|y^k - x\|^2 - \|y^k - z^k\|^2 - \|z^k - x\|^2.$$

So we have

$$
\begin{aligned}
&\|z^k - x\|^2 \\
&\leq \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\delta\|y^k - x^k\|^2 + 2\delta\|z^k - y^k\|^2 \\
&\quad -2\gamma_k[(f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x)] \\
&= \|x^k - x\|^2 - (1-2\delta)\|y^k - x^k\|^2 - (1-2\delta)\|z^k - y^k\|^2 \\
&\quad -2\gamma_k[(f+g)(y^k) - (f+g)(x) + (f+g)(z^k) - (f+g)(x)]. \qquad (4.2.35)
\end{aligned}
$$

Let $x_* = P_{S_*}F(x_*)$. Then we have

$$\|z^k - x_*\|^2 \leq \|x^k - x_*\|^2 - (1-2\delta)(\|y^k - x^k\|^2 + \|z^k - y^k\|^2). \qquad (4.2.36)$$

Now, we will show that $(x^k)_{k\in\mathbb{N}}$ is bounded. Similar to Theorem 4.2.3, we can show that

$$\|x^{k+1} - x_*\| \leq (1 - a_k(1-c))\|x^k - x_*\| + a_k\|F(x_*) - x_*\|. \tag{4.2.37}$$

By induction, we can show that $(x^k)_{k\in\mathbb{N}}$ is bounded. We also can show that

$$\begin{aligned}\|x^{k+1} - x_*\|^2 \leq\ & \frac{a_k c}{2}\|x^k - x_*\|^2 + \frac{a_k}{2}\|x^{k+1} - x_*\|^2 + a_k\langle F(x_*) - x_*, x^{k+1} - x_*\rangle \\ & + \frac{(1-a_k)}{2}(\|z^k - x_*\|^2 + \|x^{k+1} - x_*\|^2). \end{aligned} \tag{4.2.38}$$

Using (4.2.36) and (4.2.38), we then have

$$\begin{aligned} & \|x^{k+1} - x_*\|^2 \\ \leq\ & (1 - a_k(1-c))\|x^k - x_*\|^2 - (1-2\delta)(1-a_k)(\|y^k - x^k\|^2 + \|z^k - y^k\|^2) \\ & + 2a_k\langle F(x_*) - x_*, x^{k+1} - x_*\rangle. \end{aligned} \tag{4.2.39}$$

Applying Lemma 3.1.45, we set

$$\begin{aligned} s^k &= \|x^k - x_*\|^2 \\ \varphi_k &= 2a_k\langle F(x_*) - x_*, x^{k+1} - x_*\rangle \\ \chi_k &= \frac{2}{(1-c)}\langle F(x_*) - x_*, x^{k+1} - x_*\rangle \\ \psi_k &= (1-2\delta)(1-a_k)(\|y^k - x^k\|^2 + \|z^k - y^k\|^2) \\ \omega_k &= a_k(1-c). \end{aligned}$$

Let (k_n) be a subsequence of (k) and suppose that $\lim_{n\to\infty} \psi_{k_n} = 0$. Then we have $\|x^{k_n} - y^{k_n}\| \to 0$ and $\|y^{k_n} - z^{k_n}\| \to 0$ as $n \to \infty$. Also we obtain,

$$\begin{aligned} \|x^{k_n+1} - z^{k_n}\| &= \|a^{k_n}F(x^{k_n}) + (1 - a^{k_n})z^{k_n} - z^{k_n}\| \\ &= a^{k_n}\|F(x^{k_n}) - z^{k_n}\| \to 0, \text{as } n \to \infty. \end{aligned}$$

This shows that $\|x^{k_n+1} - x^{k_n}\| \to 0$ as $n \to \infty$. Since $(x^{k_n})_{n\in\mathbb{N}}$ is bounded, the set of its weak accumulation points is nonempty. Take any weak accumulation point $\bar{x}$ of $(x^{k_n})_{n\in\mathbb{N}}$. So there is a subsequence $(x^{k_{n_i}})_{i\in\mathbb{N}}$ of $(x^{k_n})_{n\in\mathbb{N}}$ weakly converging to $\bar{x}$. From Assumption (A2), we get

$$\lim_{i\to\infty} \|\nabla f(y^{k_{n_i}}) - \nabla f(z^{k_{n_i}})\| = 0. \tag{4.2.40}$$

Using $y^{k_{n_i}} = \text{prox}_{\gamma_{k_{n_i}} g}(x^{k_{n_i}} - \gamma_{k_{n_i}} \nabla f(x^{k_{n_i}}))$ and (3.1.24), we see that

$$\frac{y^{k_{n_i}} - z^{k_{n_i}} - \gamma_{k_{n_i}} \nabla f(y^{k_{n_i}})}{\gamma_{k_{n_i}}} \in \partial g(z^{k_{n_i}}) \tag{4.2.41}$$

which implies that

$$\begin{aligned} \frac{y^{k_{n_i}} - z^{k_{n_i}}}{\gamma_{k_{n_i}}} - \nabla f(y^{k_{n_i}}) + \nabla f(z^{k_{n_i}}) &\in \partial g(z^{k_{n_i}}) + \nabla f(z^{k_{n_i}}) \\ &\subseteq \partial(f+g)(z^{k_{n_i}}). \end{aligned} \tag{4.2.42}$$

Hence $\bar{x} \in S_*$ by Lemma 3.1.51. So we obtain

$$\begin{aligned} \limsup_{n\to\infty} \langle F(x_*) - x_*, x^{k_n} - x_* \rangle &= \lim_{i\to\infty} \langle F(x_*) - x_*, x^{k_{n_i}} - x_* \rangle \\ &= \langle F(x_*) - x_*, \bar{x} - x_* \rangle \\ &\leq 0. \end{aligned} \tag{4.2.43}$$

It follows that

$$\limsup_{n\to\infty} \langle F(x_*) - x_*, x^{k_n+1} - x_* \rangle \leq 0. \tag{4.2.44}$$

Hence $\limsup_{n\to\infty} \chi_{k_n} \leq 0$. Then the sequence $(x^k)_{k\in\mathbb{N}}$ converges strongly to $x_* = P_{S_*} F(x_*)$ by Lemma 3.1.45. This completes the proof. □

Proposition 4.2.8 *Let $(\gamma_k)_{k\in\mathbb{N}}$ be the sequence generated by Linesearch 4.2.4. If the gradient of f is globally Lipschitz continuous on* dom$g$ *with constant $L > 0$,* *then* $\sigma \geq \gamma_k \geq \min\{\sigma, \delta\theta/2L\}$ *for all* $k \in \mathbb{N}$.

Proof. Suppose that ∇f is globally Lipschitz continuous with constant $L > 0$. It is obvious that $\gamma_k \leq \sigma$. If $\gamma_k < \sigma$, define $\bar{\gamma}_k := \frac{\gamma_k}{\theta}$, $\bar{y}^k := \text{prox}_{\bar{\gamma}_k g}(x^k - \bar{\gamma}_k \nabla f(x^k))$ and $\bar{z}^k := \text{prox}_{\bar{\gamma}_k g}(\bar{y}^k - \bar{\gamma}_k \nabla f(\bar{y}^k))$. It follows from the definition of Linesearch 4.2.4 that

$$\begin{aligned} & 2\bar{\gamma}_k \max\{\|\nabla f(\bar{z}^k) - \nabla f(\bar{y}^k)\|, \|\nabla f(x^k) - \nabla f(\bar{y}^k)\|\} \\ > \ & \delta(\|\bar{z}^k - \bar{y}^k\| + \|x^k - \bar{y}^k\|), \end{aligned} \tag{4.2.45}$$

which gives $\|\bar{z}^k - \bar{y}^k\| + \|x^k - \bar{y}^k\| \neq 0$ for all $k \in \mathbb{N}$. By the Lipschitz assumption on ∇f, we obtain

$$\|\nabla f(\bar{z}^k) - \nabla f(\bar{y}^k)\| \leq L\|\bar{z}^k - \bar{y}^k\|$$

and

$$\|\nabla f(x^k) - \nabla f(\bar{y}^k)\| \leq L\|x^k - \bar{y}^k\|.$$

This shows that

$$\begin{aligned} & \max\{\|\nabla f(\bar{z}^k) - \nabla f(\bar{y}^k)\|, \|\nabla f(x^k) - \nabla f(\bar{y}^k)\|\} \\ \leq \ & \|\nabla f(\bar{z}^k) - \nabla f(\bar{y}^k)\| + \|\nabla f(x^k) - \nabla f(\bar{y}^k)\| \\ \leq \ & L(\|\bar{z}^k - \bar{y}^k\| + \|x^k - \bar{y}^k\|), \forall k \in \mathbb{N}. \end{aligned}$$

Combining the latter inequality with (4.2.45), we have $2\bar{\gamma}_k L > \delta$, i.e., $\gamma_k > \dfrac{\delta\theta}{2L}$ when $\gamma_k < \sigma$. □

Remark 4.2.9 Since the second part of (A2) holds even if the gradient of f is Lipschitz continuous, using Proposition 4.2.8, it follows that the stepsize γ_k imposed on Theorem 4.2.7 is also satisfied.

4.3 Numerical examples and applications

In this section, we present some numerical examples to the signal recovery. We consider our first algorithm defined by projection method and provide a comparison among Algorithm 2.1.2, Algorithm 2.1.4 and Algorithm 4.1.1. In this case, we set $Tx^k = \text{prox}_{\alpha g}(x^k - \alpha \nabla f(x^k))$. It is known that T is a nonexpansive mapping when $\alpha \in (0, \frac{2}{L})$ and L is the Lipschitz constant of ∇f. Compressed sensing can be modeled as the following underdeterminated linear equation system:

$$y = Ax + \epsilon, \tag{4.3.1}$$

where $x \in \mathbb{R}^N$ is a vector with k nonzero components to be recovered, $y \in \mathbb{R}^M$ is the observed or measured data with noisy ϵ, and $A : \mathbb{R}^N \to \mathbb{R}^M (M < N)$ is a bounded linear operator. It is known that to solve (4.3.1) can be seen as solving the LASSO problem:

$$\min_{x \in \mathbb{R}^N} \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_1, \tag{4.3.2}$$

where $\lambda > 0$. So we can apply our method for solving (4.3.2) in case $f(x) = \frac{1}{2}\|y - Ax\|_2^2$ and $g(x) = \lambda\|x\|_1$. It is noted that $\nabla f(x) = A^T(Ax - y)$.

In our experiment, the sparse vector $x \in \mathbb{R}^N$ is generated from uniform distribution in the interval [-2,2] with k nonzero elements. The matrix $A \in \mathbb{R}^{M \times N}$ is generated from a normal distribution with mean zero and invariance one. The observation y is generated by with Gaussian noise white signal - to - noise ratio SNR=40. The initial point x^0 is picked randomly. The restoration accuracy is measured by the mean squared error as follows:

$$\text{MSE} = \frac{1}{N}\|x^k - x^*\|_2^2 < 10^{-5},$$

where x^* is an estimated signal of x.

In what follows, let $\sigma = 5$, $\theta = 0.4$, and $\delta = 0.4$ in both Algorithm 2.1.4 and Algorithm 4.1.1 and let the step size α_k in Algorithm 2.1.2 and Algorithm 2.1.4 be $\frac{1}{\|A\|^2}$. Let $h(x) = \frac{x}{5}$ be a contraction and choose $a_k = \frac{1}{100k}$ in all algorithms. We denote by CPU the time using in CPU and Iter the number of iterations. The numerical results are reported as follows:

Table 1: Computational results for solving the LASSO problem by Algorithm 4.1.1, Algorithm 2.1.4 and Algorithm 2.1.1

m - sparse signal	Method	$N = 512$, $M = 256$		$N = 1024$, $M = 512$	
		CPU	Iter	CPU	Iter
$m = 20$	Algorithm 4.1.1	4.3612	673	35.2779	1258
	Algorithm 2.1.4	41.5479	3645	265.4392	6851
	Algorithm 2.1.1	9.7712	1742	65.0949	3249
$m = 30$	Algorithm 4.1.1	6.0680	793	32.7622	1335
	Algorithm 2.1.4	56.6697	4370	282.0070	7265
	Algorithm 2.1.1	13.0234	2109	64.8357	3457
$m = 40$	Algorithm 4.1.1	5.5765	790	35.2468	1391
	Algorithm 2.1.4	57.1358	4495	324.6561	7639
	Algorithm 2.1.1	14.2279	2175	71.0742	3649
$m = 50$	Algorithm 4.1.1	7.8385	1024	41.1793	1416
	Algorithm 2.1.4	96.3842	5901	357.4149	7818
	Algorithm 2.1.1	24.7290	2873	88.7461	3731

The data in Table 1 shows that, for a given tolerance, all algorithms can be used to solve the LASSO problem in compressed sensing. To be more precise, Algorithm 4.1.1 with a linesearch take significantly less number of iterations and CPU time compared to Algorithm 2.1.1 of [18] and Algorithm 2.1.4 of [3]. Next, we provide some numerical experiments for two cases to illustrate the convergence behavior of all algorithm in comparison.

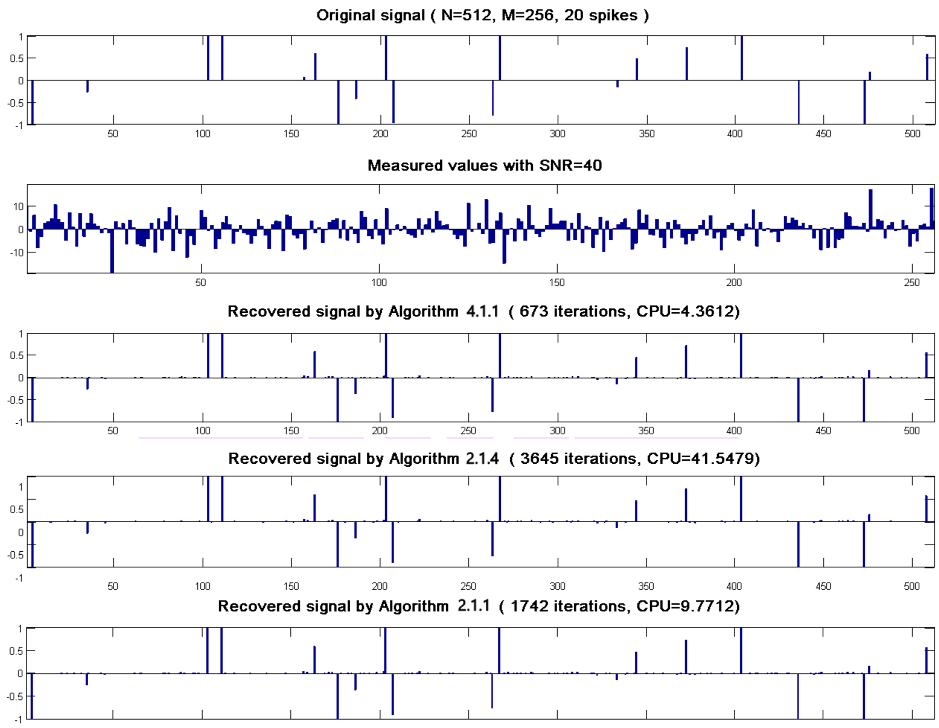

Figure 1: From top to bottom: original signal, observation data, recovered signal by Algorithm 4.1.1, Algorithm 2.1.4 and Algorithm 2.1.1 with $N = 512$ and $M = 256$, respectively.

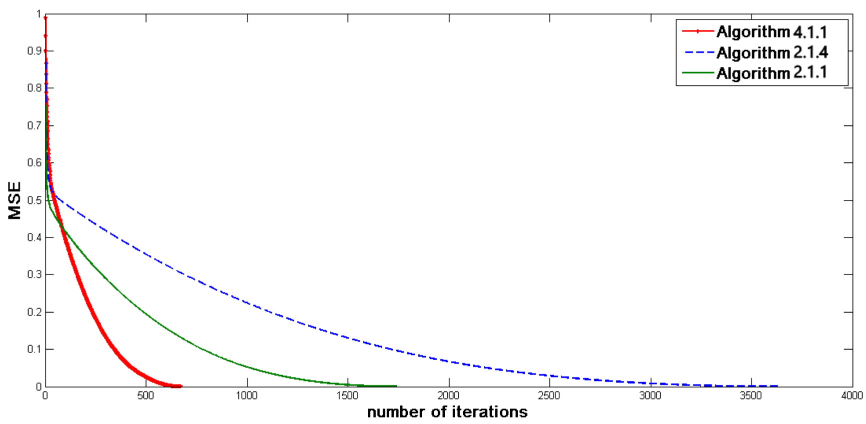

Figure 2: The MSE versus number of iterations in case N = 512, M = 256.

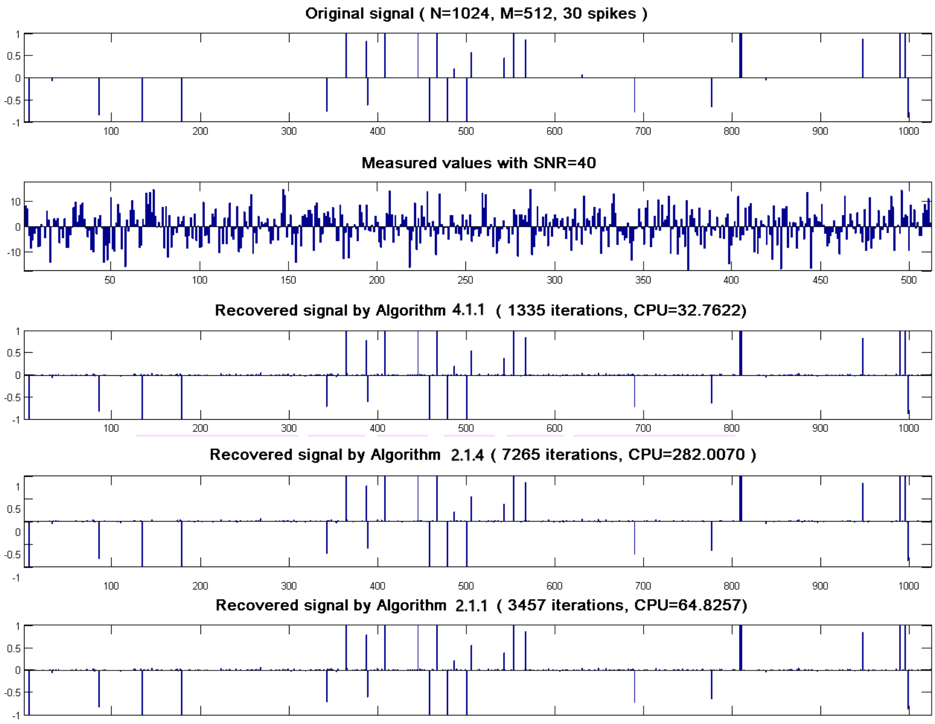

Figure 3: From top to bottom: original signal, observation data, recovered signal by Algorithm 4.1.1, Algorithm 2.1.4 and Algorithm 2.1.1 with $N = 512$ and $M = 256$, respectively.

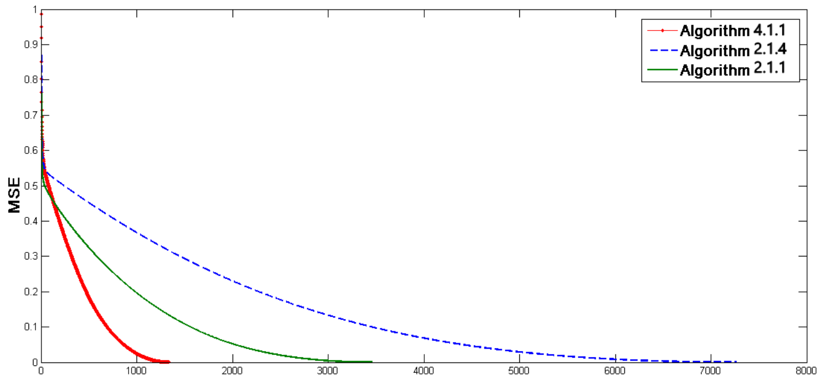

Figure 4: The MSE versus number of iterations in case N = 512, M = 256.

Next, we discuss our forward-backward algorithm defined by the shrinking projection method. We provide a comparison among Algorithm 2.1.1, Algorithm 2.1.5 and Algorithm 4.1.3. For convenience, we set all condition as in the previous example.

Table 2: Computational results for solving the LASSO problem by Algorithm 4.1.3, Algorithm 2.1.5 and Algorithm 2.1.1

m - sparse signal	Method	$N = 512$, $M = 256$		$N = 1024$, $M = 512$	
		CPU	Iter	CPU	Iter
$m = 20$	Algorithm 4.1.3	5.2158	660	40.9654	1247
	Algorithm 2.1.5	29.0458	3548	183.4528	6789
	Algorithm 2.1.1	5.8716	1696	46.3793	3222
$m = 30$	Algorithm 4.1.3	8.7975	865	45.0510	1369
	Algorithm 2.1.5	42.1279	4820	236.4308	7648
	Algorithm 2.1.1	9.8976	2325	62.7609	3645
$m = 40$	Algorithm 4.1.3	7.4329	926	42.7707	1365
	Algorithm 2.1.5	53.7019	5079	224.0050	7551
	Algorithm 2.1.1	12.2709	2461	54.1403	3608
$m = 50$	Algorithm 4.1.3	8.6868	1099	56.2084	1508
	Algorithm 2.1.5	107.2563	6309	308.9143	8439
	Algorithm 2.1.1	20.1471	3085	67.8876	4053

The data in Table 2 shows that, for a given tolerance, all algorithms can be used to solve the LASSO problem in compressed sensing. To be more precise, Algorithm 4.1.3 with a linesearch take significantly less number of iterations and CPU time compared to Algorithm 2.1.1 of [18] and Algorithm 2.1.5 of [23].

We plot the original signal, observation data, recovered signal, the number of iterations versus MSE.

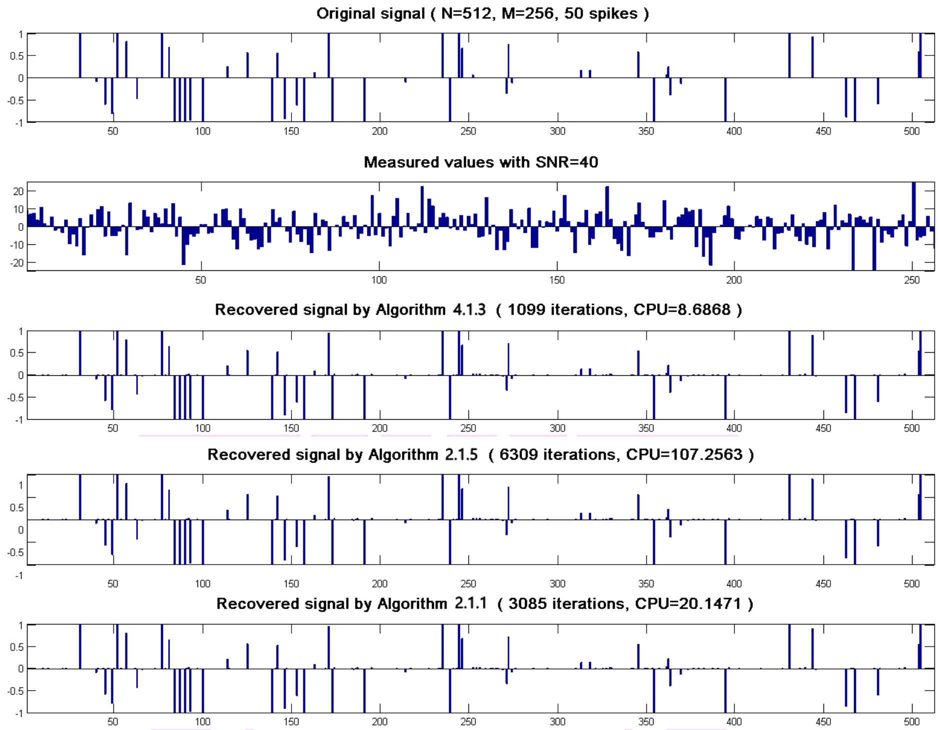

Figure 5: From top to bottom: original signal, observation data, recovered signal by Algorithm 4.1.3, Algorithm 2.1.5 and Algorithm 2.1.1 with $N = 512$ and $M = 256$, respectively.

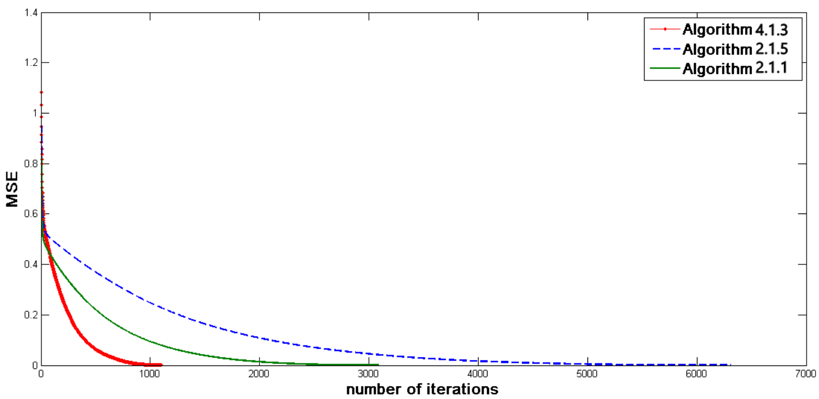

Figure 6: The MSE versus number of iterations in case N = 512, M = 256.

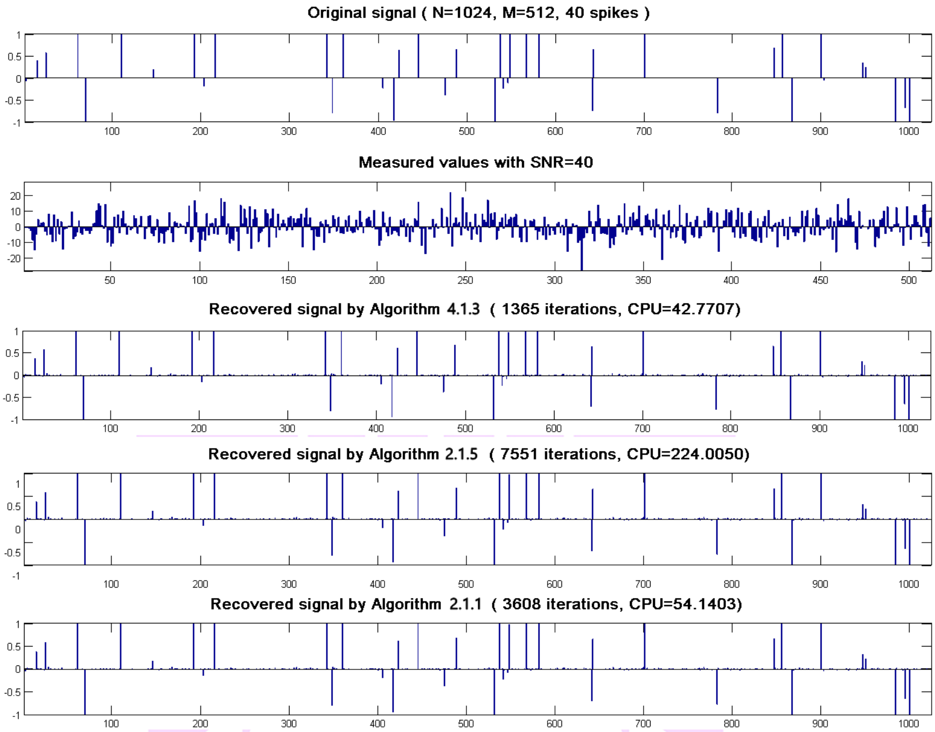

Figure 7: From top to bottom: original signal, observation data, recovered signal by Algorithm 4.1.3, Algorithm 2.1.5 and Algorithm 2.1.1 with $N = 512$ and $M = 256$, respectively.

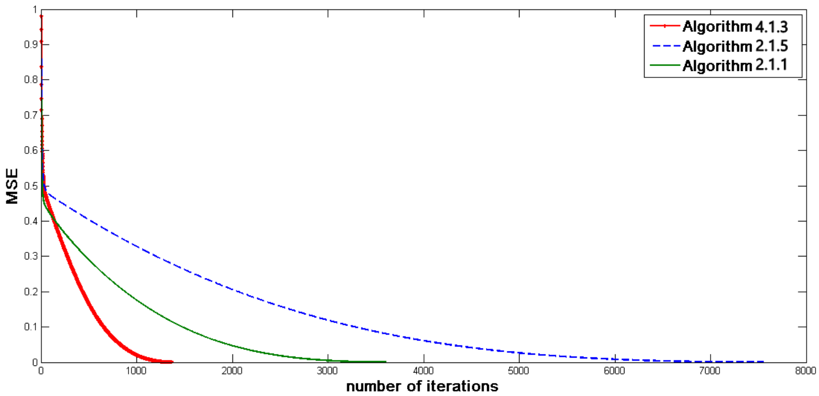

Figure 8: The MSE versus number of iterations in case N = 512, M = 256.

Next, we provide a comparison among Algorithm 2.1.2, Algorithm 4.2.1 and Algorithm 4.2.6. In our experiment, the sparse vector $x \in \mathbb{R}^N$ is generated from uniform distribution in the interval [-2,2] with k nonzero elements. The matrix $A \in \mathbb{R}^{M \times N}$ is generated from a normal distribution with mean zero and invariance one. The observation y is generated by with Gaussian noise white signal-to-noise ratio SNR=40. The initial point x^0 is picked randomly. The restoration accuracy is measured by the mean squared error as follows:

$$\text{MSE} = \frac{1}{N}\|x^k - x^*\|^2 < 10^{-5},$$

where x^* is an estimated signal of x.

In what follows, let the step size β_k in Algorithm 2.1.2 is $\frac{1}{\|A\|^2}$ and $b^k = 0$ and let $\sigma = 2,\ \theta = 0.4,$ and $\delta = 0.4$ in both Algorithm 4.2.1 and Algorithm 4.2.6. Let $F(x) = \frac{x}{2}$ be a contraction and choose $a_k = \frac{1}{100k}$ in all algorithms. We denote by CPU the time using in CPU and Iter the number of iterations. The numerical results are reported as follows:

Table 3: Computational results for solving the LASSO problem by Algorithm 2.1.2, Algorithm 4.2.1 and Algorithm 4.2.6

m - sparse signal	Method	$N = 512$, $M = 256$		$N = 1024$, $M = 512$	
		CPU	Iter	CPU	Iter
$m = 10$	Algorithm 2.1.2	0.8440	966	8.9773	2019
	Algorithm 4.2.1	0.4650	338	3.0526	635
	Algorithm 4.2.6	0.3661	180	2.7670	349
$m = 15$	Algorithm 2.1.2	1.1808	1071	8.3444	1895
	Algorithm 4.2.1	0.5427	369	3.0859	601
	Algorithm 4.2.6	0.4676	193	2.4348	324
$m = 20$	Algorithm 2.1.2	1.3669	1243	11.3604	2301
	Algorithm 4.2.1	0.6102	426	3.4292	683
	Algorithm 4.2.6	0.4997	224	2.8893	381
$m = 30$	Algorithm 2.1.2	2.8177	1753	14.5687	2587
	Algorithm 4.2.1	0.8495	572	3.7913	762
	Algorithm 4.2.6	0.7782	316	3.2678	462

The data in Table 1 shows that for a given tolerance, all algorithms can be used to solve the LASSO problem in compressed sensing. To be more precise, Algorithm 4.2.6 with a new linesearch take significantly less number of iterations and CPU time compared to Algorithm 2.1.2 of [31] and Algorithm 4.2.1 with Linesearch 2.1.4.

We next give some numerical experiments for two cases to illustrate the convergence behavior of all algorithm in comparison. We plot the original signal, observation data, recovered signal, the number of iterations versus objective function value and MSE.

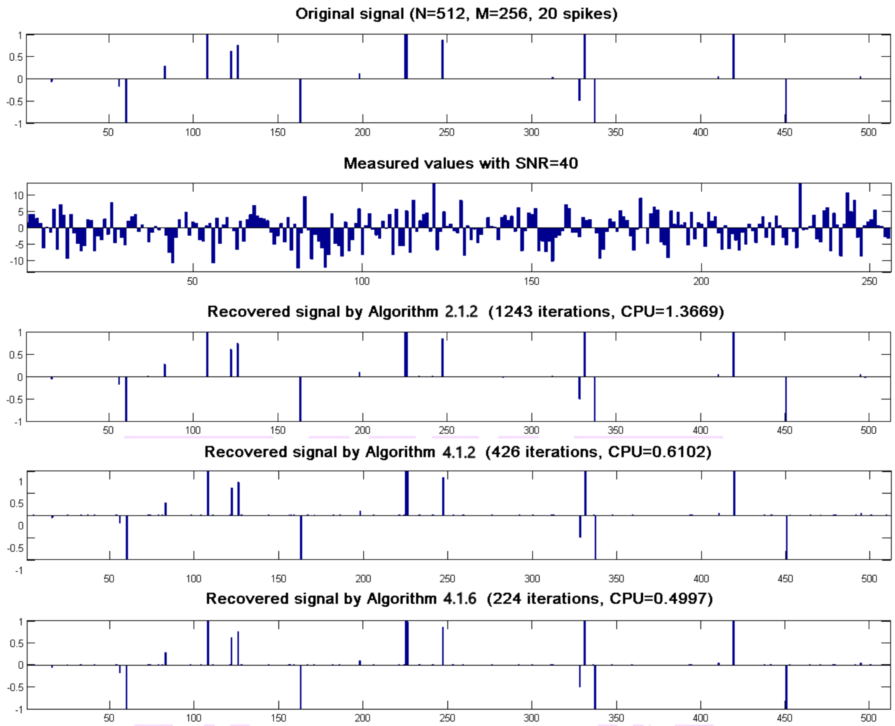

Figure 9: From top to bottom: original signal, observation data, recovered signal by Algorithm 2.1.2, Algorithm 4.2.1 and Algorithm 4.2.6 with $N = 512$ and $M = 256$, respectively.

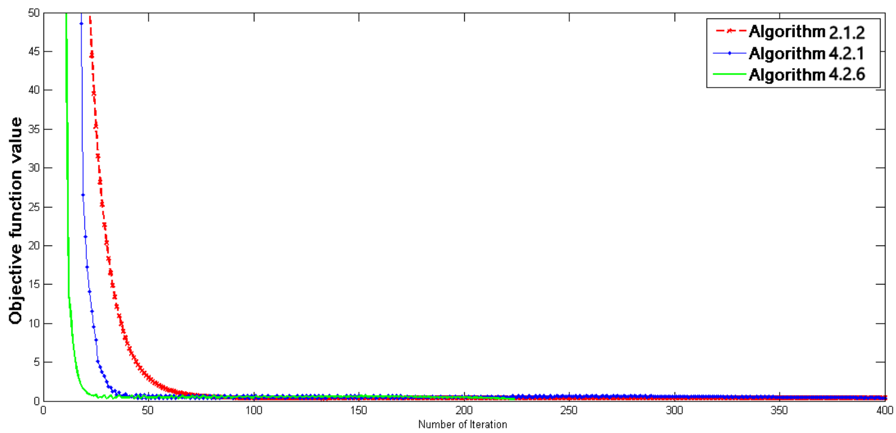

Figure 10: The objective function value versus number of iterations in case N = 512, M = 256.

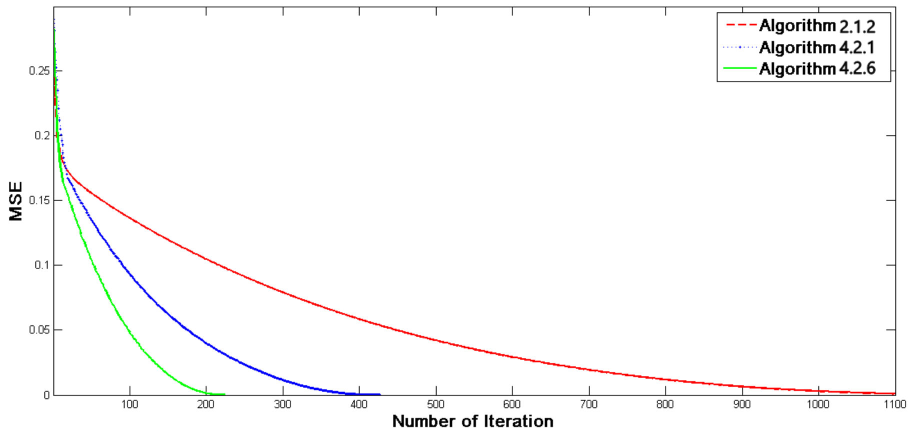

Figure 11: The MSE versus number of iterations in case N = 512, M = 256.

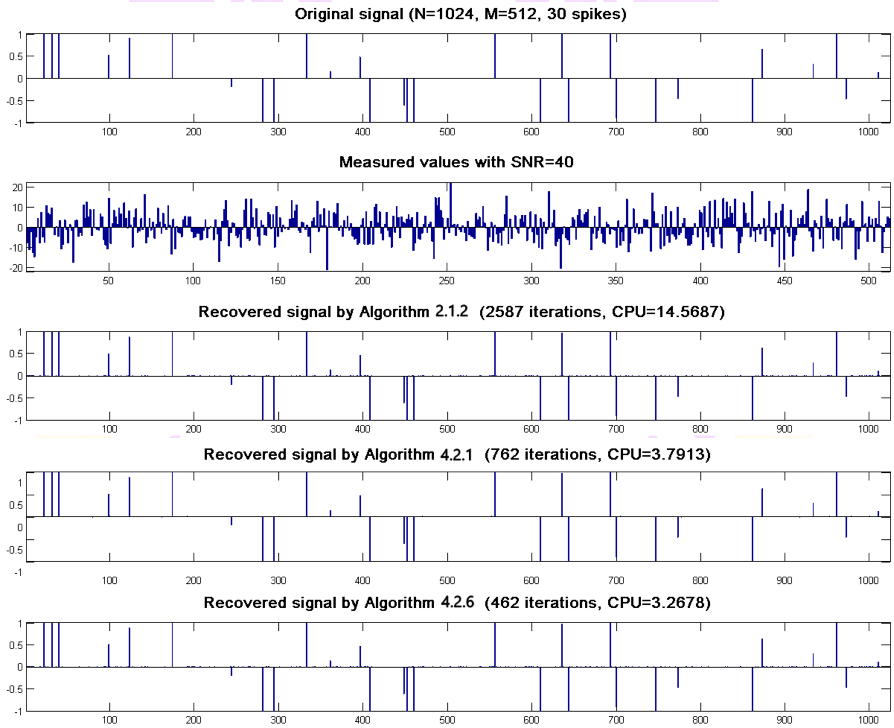

Figure 12: From top to bottom: original signal, observation data, recovered signal by Algorithm 2.1.2, Algorithm 4.2.1 and Algorithm 4.2.6 with $N = 1024$ and $M = 512$, respectively.

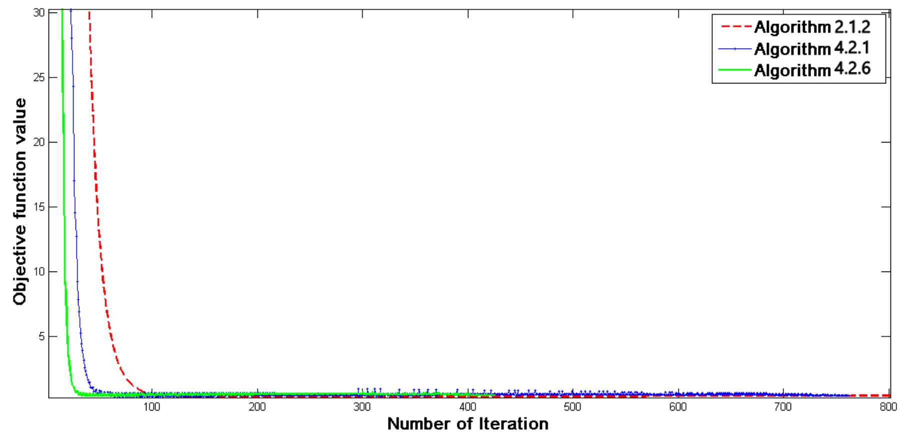

Figure 13: The objective function value versus number of iterations in case N = 1024, M = 512.

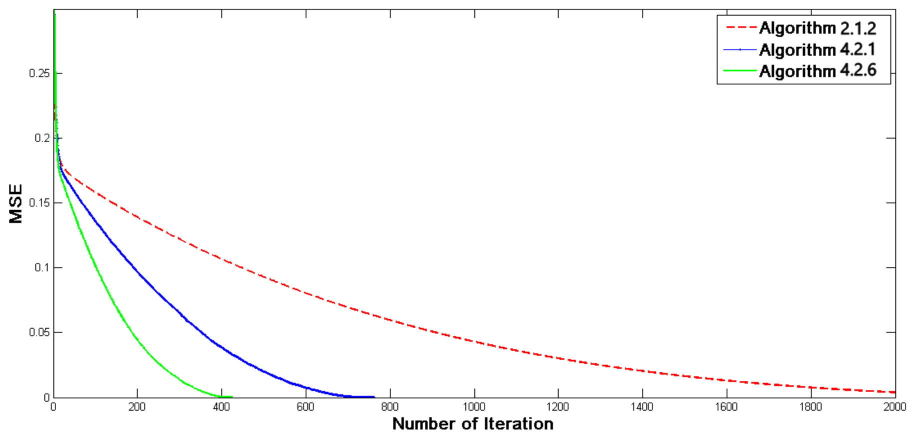

Figure 14: The MSE versus number of iterations in case N = 1024, M = 512.

CHAPTER V

CONCLUSIONS

From our study, we get the main results as the following:

Algorithm 5.3.1 *(step 0) Choose* $x^0 \in \mathrm{dom}g$*, take* $\delta \in (0, \frac{1}{2})$*,* $\sigma > 0$ *and* $\theta \in (0, 1)$.

(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\begin{aligned} & \alpha_k \|\nabla f(\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k))) - \nabla f(x^k)\| \\ \leq \ & \delta \|\mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)) - x^k\|. \end{aligned} \tag{5.3.3}$$

(step 2) Set

$$y^k = \mathrm{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{5.3.4}$$

(step 3) Compute

$$C_k = \{x_* \in \mathrm{dom}g : \|y^k - x_*\| \leq \|x^k - x_*\|\}$$

and

$$Q_k = \{x_* \in \mathrm{dom}g : \langle x_* - x^k, x^0 - x^k \rangle \leq 0\}. \tag{5.3.5}$$

(step 4) Compute

$$x^{k+1} = P_{C_k \cap Q_k}(x^0). \tag{5.3.6}$$

(step 5) Set $k \leftarrow k + 1$*, and go to (step 1).*

Theorem 5.3.2 *Let* H *be a real Hilbert space. Assume that there exists* $\alpha > 0$ *such that* $\alpha_k \geq \alpha > 0$*. Then the sequence* $(x^k)_{k=0}^{\infty}$ *generated by Algorithm 4.1.1 converges strongly to* $\bar{x} = P_{S_*}(x^0)$.

Algorithm 5.3.3 *(step 0) Set* $C_0 = \mathrm{dom}g$*, choose* $x^0 \in \mathrm{dom}g$*, take* $\delta \in (0, \frac{1}{2})$*,* $\sigma > 0$ *and* $\theta \in (0, 1)$.

(step 1) Set $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k\|\nabla f(y^k) - \nabla f(x^k)\| \le \delta\|y^k - x^k\|. \tag{5.3.7}$$

(step 2) Set

$$y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)). \tag{5.3.8}$$

(step 3) Compute

$$C_{k+1} = \{x_* \in C_k : \|y^k - x_*\| \le \|x^k - x_*\|\}. \tag{5.3.9}$$

(step 4) Compute

$$x^{k+1} = P_{C_{k+1}}(x^0). \tag{5.3.10}$$

(step 5) Set $k \leftarrow k+1$*, and go to (step 1).*

Theorem 5.3.4 *Let* H *be a real Hilbert space. Assume that there exists* $\alpha > 0$ *such that* $\alpha_k \ge \alpha > 0$*. Then the sequence* $(x^k)_{k=0}^{\infty}$ *generated by Algorithm 4.1.3 converges strongly to* $\bar{x} = P_{S_*}(x^0)$*.*

Algorithm 5.3.5 *Let* $F : \text{domg} \to \text{domg}$ *be a contraction. Let* $\sigma > 0$*,* $\theta \in (0,1)$ *and* $\delta \in (0, \frac{1}{2})$*, take* $x^0 \in$ domg *and*

$$y^k = \text{prox}_{\alpha_k g}(x^k - \alpha_k \nabla f(x^k)) \tag{5.3.11}$$

where $\alpha_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\alpha_k\|\nabla f(y^k) - \nabla f(x^k)\| \le \delta\|y^k - x^k\|. \tag{5.3.12}$$

Construct x^{k+1} *by*

$$x^{k+1} = a_k F(x^k) + (1 - a_k)y^k. \tag{5.3.13}$$

Lemma 5.3.6 *[3] The linesearch (4.2.2) stops after finitely many steps.*

Theorem 5.3.7 *Let $(x^k)_{k\in\mathbb{N}}$ and (α_k) be sequences generated by Algorithm 4.2.1. Suppose that there exists $\alpha > 0$ such that $\alpha_k \geq \alpha$ for all $k \in \mathbb{N}$ and $(a_k) \subset (0,1)$ such that*

$$\lim_{k\to\infty} a_k = 0 \text{ and } \sum_{k=1}^{\infty} a_k = \infty,$$

then the sequence $(x^k)_{k\in\mathbb{N}}$ converges strongly to a point $x_* = P_{S_*}F(x_*)$.

Linesearch 5.3.8 *Given $x \in \operatorname{dom} g$, $\sigma > 0$, $\theta \in (0,1)$ and $\delta \in (0, \frac{1}{2})$.*

For $i = 0, 1, 2, \dots$, set

$$L(x,i) = \operatorname{prox}_{\sigma\theta^i g}(x - \sigma\theta^i \nabla f(x))$$

and

$$S(x,i) = \operatorname{prox}_{\sigma\theta^i g}(L(x,i) - \sigma\theta^i \nabla f(L(x,i))).$$

If

$$\begin{aligned} & 2\sigma\theta^i \max\{\|\nabla f(S(x,i)) - \nabla f(L(x,i))\|, \|\nabla f(L(x,i)) - \nabla f(x)\|\} \\ \leq \ & \delta(\|S(x,i) - L(x,i)\| + \|L(x,i) - x\|) \end{aligned} \quad (5.3.14)$$

then $\gamma = \sigma\theta^i$

Else $i = i + 1$.

Linesearch 5.3.8 is well - defined.

Lemma 5.3.9 *The Linesearch 5.3.8 stops after finitely many steps.*

Using this linesearch, we propose the following algorithm:

Algorithm 5.3.10 *Let $F : \operatorname{dom} g \to \operatorname{dom} g$ be a contraction. Let $\sigma > 0$, $\theta \in (0,1)$*

and $\delta \in (0, \frac{1}{2})$, *take* $x^0 \in \operatorname{dom} g$ *and*

$$y^k = \operatorname{prox}_{\gamma_k g}(x^k - \gamma_k \nabla f(x^k)), \tag{5.3.15}$$

where $\gamma_k = \sigma\theta^{m_k}$ *and* m_k *is the smallest nonnegative integer such that*

$$\begin{aligned} & 2\gamma_k \max\{\|\nabla f(\operatorname{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k))) - \nabla f(y^k)\|, \|\nabla f(x^k) - \nabla f(y^k)\|\} \\ \leq\ & \delta(\|\operatorname{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k)) - y^k\| + \|x^k - y^k\|). \end{aligned} \tag{5.3.16}$$

Construct x^{k+1} *by*

$$x^{k+1} = a_k F(x^k) + (1 - a_k)\operatorname{prox}_{\gamma_k g}(y^k - \gamma_k \nabla f(y^k)). \tag{5.3.17}$$

Theorem 5.3.11 *Let* $(x^k)_{k\in\mathbb{N}}$ *and* (γ_k) *be sequences generated by Algorithm 4.2.6. Suppose that there exists* $\gamma > 0$ *such that* $\gamma_k \geq \gamma$ *for all* $k \in \mathbb{N}$ *and* $(a_k) \subset (0, 1)$ *such that*

$$\lim_{k\to\infty} a_k = 0 \text{ and } \sum_{k=1}^{\infty} a_k = \infty,$$

then the sequence $(x^k)_{k\in\mathbb{N}}$ *converges strongly to a point* $x_* = P_{S_*}F(x_*)$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Bauschke, H. H., and Combettes, P. L. (2001). A weak to strong convergence principle for Fejér - monotone methods in Hilbert spaces. **Mathematics of operations research**, 26(2), 248 - 264.

[2] Bauschke, H. H. and Combettes, P. L. (2011). Convex Analysis and Monotone Operator Theory in Hilbert Spaces. New York: **Springer**.

[3] Bello Cruz, J. Y. and Nghia, T. T. (2016). On the convergence of the forward-backward splitting method with linesearches. **Optimization Methods and Software**, 31(6), 1209 - 1238.

[4] Burachik, R. S. and Iusem, A. N. (2007). Set - valued mappings and enlargements of monotone operators (Vol. 8). **Springer Science Business Media**.

[5] Chen, G. H., and Rockafellar, R. T. (1997). Convergence rates in forward-backward splitting. **SIAM Journal on Optimization**, 7(2), 421 - 444.

[6] Cholamjiak, P. (2016). A generalized forward-backward splitting method for solving quasi inclusion problems in Banach spaces. **Numerical Algorithms**, 71(4), 915 - 932.

[7] Cholamjiak, W., Cholamjiak, P., and Suantai, S. (2018). An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces. **Journal of Fixed Point Theory and Applications**, 20(1), 42.

[8] Combettes, P. L. and Pesquet, J. C. (2011). Proximal splitting methods in signal processing. In Fixed-point algorithms for inverse problems in science and engineering (185 - 212). **Springer**, New York, NY.

[9] Combettes, P. L. and Wajs, V. R. (2005). Signal recovery by proximal forward-backward splitting. **Multiscale Modeling Simulation**, 4(4), 1168 - 1200.

[10] Dong, Q., Jiang, D., Cholamjiak, P., and Shehu, Y. (2017). A strong convergence result involving an inertial forward-backward algorithm for monotone inclusions. **Journal of Fixed Point Theory and Applications**, 19(4), 3097 - 3118.

[11] Dunn, J. C. (1976). Convexity, monotonicity, and gradient processes in Hilbert space. **Journal of Mathematical Analysis and Applications**, 53(1), 145 - 158.

[12] Guler, O. (1991). On the convergence of the proximal point algorithm for convex minimization. **SIAM Journal on Control and Optimization**, 29(2), 403 - 419.

[13] He, S., and Yang, C. (2013). Solving the variational inequality problem defined on intersection of finite level sets. **Abstract and Applied Analysis**, Volume 2013, Article ID 942315, 8 pages, http://dx.doi.org/10.1155/2013/942315.

[14] Iusem, A. N., Svaiter, B. F., and Teboulle, M. (1994). Entropy - like proximal methods in convex programming. **Mathematics of Operations Research**, 19(4), 790 - 814.

[15] Kankam, K., Pholasa, N., and Cholamjiak, P. On convergence and complexity of the modified forward-backward method involving new linesearches for convex minimization. **Mathematical Methods in the Applied Sciences**. (https://doi.org/10.1002/mma.5420)

[16] Khan, S. A., Suantai, S., and Cholamjiak, W. (2018). Shrinking projection methods involving inertial forward-backward splitting methods for inclusion problems. Revista de la Real Academia de Ciencias Exactas, Fasicasy Naturales. Serie A. **Matematicas**, 1 - 12.

[17] S. A., Cholamjiak, W., and Kazmi, K. R. (2018). An inertial forward-backward splitting method for solving combination of equilibrium problems and inclusion problems. **Computational and Applied**

Mathematics, 37(5), 6283 - 6307.

[18] Lin, L. J., and Takahashi, W. (2012). A general iterative method for hierarchical variational inequality problems in Hilbert spaces and applications. **Positivity**, 16(3), 429 - 453.

[19] Lions, P. L., and Mercier, B. (1979). Splitting algorithms for the sum of two nonlinear operators. **SIAM Journal on Numerical Analysis**, 16(6), 964 - 979.

[20] Lorenz, D. A., and Pock, T. (2015). An inertial forward-backward algorithm for monotone inclusions. **Journal of Mathematical Imaging and Vision**, 51(2), 311 - 325.

[21] Martinet, B. (1970). Brvèv communication. Régularisation d'inéquations variationnelles par approximations successives. Revue franaise d'informatique et de recherche opérationnelle. **Série rouge**, 4(R3), 154 - 158.

[22] Moudafi, A. (2000). Viscosity approximation methods for fixed-points problems. **Journal of Mathematical Analysis and Applications**, 241(1), 46 - 55.

[23] Nakajo, K. and Takahashi, W. (2003). Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups. **Journal of Mathematical Analysis and Applications**, 279(2), 372 - 379.

[24] Parikh, N. and Boyd, S. (2014). Proximal algorithms. **Foundations and Trends in Optimization**, 1(3), 127 - 239.

[25] Passty, G. B. (1979). Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. **Journal of Mathematical Analysis and Applications**, 72, 383 - 290.

[26] Pholasaa, N., Cholamjiaka, P., and Cho, Y. J. (2016). Modified forward-backward splitting methods for accretive operators in Banach spaces. **J. Nonlinear Sci. Appl**, 9, 2766 - 2778.

[27] Rockafellar, R. T. (1976). Monotone operators and the proximal point algorithm. **SIAM Journal on Control and Optimization**, 14(5), 877 - 898.

[28] Takahashi, W. and Zembayashi, K. (2008). Strong convergence theorem by a new hybrid method for equilibrium problems and relatively nonexpansive mappings. **Fixed Point Theory and Applications**, 2008(528476), 528476.

[29] Tseng, P. (2000). A modified forward-backward splitting method for maximal monotone mappings. **SIAM Journal on Control and Optimization**, 38(2), 431 - 446.

[30] Wang, C. and Xiu, N. (2000). Convergence of the gradient projection method for generalized convex minimization. **Computational Optimization and Applications**, 16(2), 111 - 120.

[31] Wang, Y., and Wang, F. (2018). Strong convergence of the forward-backward splitting method with multiple parameters in Hilbert spaces. **Optimization**, 67(4), 493 - 505.

[32] Xu, H. K. (2011). Averaged mappings and the gradient-projection algorithm. **Journal of Optimization Theory and Applications**, 150(2), 360 - 378.

BIOGRAPHY

BIOGRAPHY

Name Surname Kunrada Kankam

Date of Birth September 11, 1995

Place of Birth Phrae Province, Thailand

Address 228 Moo 7, Mae Chua Sub - district,
Den Chai District, Phrae Province,
Thailand 54110

Education Background

2020 M.Sc. (Mathematics), University of Phayao, Phayao, Thailand

2017 B.Sc. (Mathematics), University of Phayao, Phayao, Thailand

Publications

Articles

1. **Kankam, K.**, Pholasa, N., and Cholamjiak, P. (2019). On convergence and complexity of the modified forward-backward method involving new linesearches for convex minimization. Mathematical Methods in the Applied Sciences, 42(5), 1352 - 1362.

2. **Kankam, K.**, Pholasa, N., and Cholamjiak, P. (2019). Hybrid forward-backward algorithms using linesearch rule for minimization problem. Thai Journal of Mathematics, 17(3), 607 - 625.

3. Suantai, S., **Kankam, K.**, and Cholamjiak, P. (2020). A novel forward-backward algorithm for solving convex minimization problem in Hilbert spaces. Mathematics, 8(1), 42.

Others

1. **Kankam, K.**, Srinak, P., Cholamjiak, P. and Pholasa, N. (2019). Solving the multiple-set split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. Rendiconti del Circolo Matematico di Palermo Series 2, 1 - 18.

Conference presentations

1. **Kankam, K.** (May 24 - 25, 2018). Solving the multiple-sets split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. In The 10th National Science Research Conference, Mahasarakham University, Mahasarakham, Thailand.

2. **Kankam, K.** (November 5 - 9, 2018). Solving the multiple-sets split feasibility problem and the equilibrium problem by a new relaxed CQ algorithm. In The 6th Asian Conference on Nonlinear Analysis and Optimization, Okinawa, Japan.

3. **Kankam, K.** (September 12 - 14, 2019). Strong convergence of the forward-backward splitting algorithms via linesearches. In International Workshop on Applied Nonlinear Analysis, Bangsaen, Chonburi, Thailand.