1. Introduction.
2. Attouch H, Cabot A. Convergence of rates of inertial forward-backward algorithms[J]. Siam Journal on Optimization, 2018, 28(1): 849-874.
3. Wen B, Xue X. On the convergence of the iterates of proximal gradient algorithm with extrapolation for convex nonsmooth minimization problems[J]. Journal of Global Optimization, 2019, 75(3): 767-787.
4. Liu H , Wang T , Liu Z . Convergence Rate of Inertial Forward-Backward Algorithms Based on the Local Error Bound Condition[J]. 2020.arXiv:2007.07432 [math.OC]

Introduction

we consider the following problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n}\{F(x):=f(\mathbf{x})+g(\mathbf{x})\} \tag{1}$$

Assume:

(1) f is a smooth convex function and continuously differentiable with L_f-Lipschitz continuous gradient;

(2) g be a proper,convex,and lower semicontinuous;

(3) F is level bounded,$S = argminF$ is nonempty.

Introduction

IFB:

$$\begin{cases} y_k = x_k + \alpha_k \left(x_k - x_{k-1}\right) \\ x_{k+1} = \operatorname{prox}_{sg}\left(y_k - s\nabla f\left(y_k\right)\right) \end{cases}$$

General case:

$$\begin{cases} y^k = x^k + \alpha_k \left(x^k - x^{k-1}\right) \\ x^{k+1} = \underset{x\in\mathbb{R}^n}{\arg\min}\left\{\left\langle \nabla f\left(y^k\right), x\right\rangle + \frac{1}{2s}\left\|x - y^k\right\|^2 + g(x)\right\} \end{cases}$$

Where $s \in (0, \frac{1}{L}]$,α_k is nonnegative.

Introduction

Known conclusion

❶ For $t_1 = 1, t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$,$\alpha_k^{[1]} = \frac{t_k-1}{t_{k+1}}$,$F(x_k) - F(x^*) = O(\frac{1}{k^2})$;

❷ For $\alpha > 2$,$\alpha_k^{[2]} = \frac{k-1}{k+\alpha}(\alpha > 3, \alpha_k = \frac{k-1}{k+\alpha-1})$,

$F(x_k) - F(x^*) = O(\frac{1}{k^2})$,sequence $\{x_k\}$ given by FISTA converges weakly to a minimizer of F;

[1] Beck A, Teboulle M. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems[J]. Siam Journal on Imaging Sciences, 2009, 2(1): 183-202.

[2] Chambolle A, Dossal C. On the convergence of the iterates of the fast iterative shrinkage/thresholding algorithm[J]. Journal of Optimization Theory and Applications, 2015, 166(3): 968-982.

Background

Problem

$$\min\{F(x) := f(\mathbf{x}) + g(\mathbf{x}) :\ x \in \mathcal{H}\} \tag{2}$$

Assume:

(1) $f : \mathcal{H} \to \mathbb{R}$ is a smooth convex function and continuously differentiable with L_f-Lipschitz continuous gradient;

(2) $g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is proper,convex,and lower semicontinuous;

(3) $S = argminF$ is nonempty.

Introduction

Known conclusion

❸ In $\mathcal{H}$ for $\alpha > 2$,$\alpha_k^{[3]} = \frac{k-1}{k+\alpha}$($\alpha > 3$, $\alpha_k = \frac{k-1}{k+\alpha-1}$),

$F(x_k) - F(x^*) = o(\frac{1}{k^2})$,sequence $\{x_k\}$ given by FISTA converges weakly to a minimizer of F.

$$\lim_{k\to\infty} k^2(F(x_k) - F(x^*)) = 0 \quad \lim_{k\to\infty} k\|x_{k+1} - x_k\| = 0. \tag{3}$$

In other words ,$F(x_k) - F(x^*) = o(\frac{1}{k^2}) \quad \|x_{k+1} - x_k\| = o(\frac{1}{k})$

[3]Attouch H, Peypouquet J. The rate of convergence of Nesterov's accelerated Forward-Backward method is actually faster than $1/k^2$[J]. Siam Journal on Optimization, 2016, 26(3): 1824-1834.

Introduction

Proof:Main idea

Let $x^* \in argminF = S$,set $\varepsilon(k) = \frac{2s}{\alpha-1}(F(x_k) - F(x^*)) + (\alpha - 1)\|z_k - x^*\|^2$.
Where $z_k = \frac{k-\alpha-1}{\alpha-1} y_k - \frac{k}{\alpha-1} x_k = x_k + \frac{k-1}{\alpha-1}(x_k - x_{k-1})$

❶ The sequence $\{\varepsilon(k)\}$ is non increasing and $\lim_{k\to\infty} \varepsilon(k)$ exists;

❷ For each $k \geq 0$,we have $F(x_k) - F(x^*) \leq \frac{(\alpha-1)\varepsilon(0)}{2s(k+\alpha-2)^2}$ and $\|z_k - x^*\|^2 \leq \frac{\varepsilon(0)}{\alpha-1}$;

❸ If $\alpha > 3$,then $\sum_{k=1}^{\infty} k(F(x_k) - F(x^*)) \leq \frac{(\alpha-1)\varepsilon(1)}{2s(\alpha-3)}$;

❹ If $\alpha > 3$,then $\sum_{k=1}^{\infty} k\|x_{k+1} - x_k\|^2 \leq \frac{(\alpha-1)\varepsilon(1)}{s(\alpha-3)}$

Lemma

If $\alpha > 3$,then $\lim_{k\to\infty}[k^2\|x_{k+1} - x_k\|^2 + (k+1)^2(F(x_{k+1}) - F(x^*))]$ exists.

❸ and ❹ imply $\sum_{k=1}^{\infty} \frac{1}{k}[k^2\|x_{k+1} - x_k\|^2 + (k+1)^2(F(x_{k+1}) - F(x^*))] < +\infty$

All the terms are nonnegative ,then both limits are 0.

Introduction

Lemma

Let S be a nonempty subset of $\mathcal{H}$ and $\{x_k\}$ a sequence of elements of $\mathcal{H}$

.Assume that

(i) every sequential weak cluster point of $\{x_k\}$,as $k \to +\infty$ belongs to S;

(ii) for every $z \in S$,$\lim_{k\to+\infty} \|x_k - z\|$ exists.

Then $\{x_k\}$ converges weakly as $k \to +\infty$ to a point in S.

1. $\lim_{k\to\infty} \|z_k - x^*\|$ exists $\Rightarrow \lim_{k\to\infty} \|x_k - x^*\|$ exists;
2. $x^* \in S$

1

- Attouch H, Cabot A. Convergence of rates of inertial forward-backward algorithms[J]. Siam Journal on Optimization, 2018, 28(1): 849-874.

Background

Problem

$$\min\{F(x) := f(\mathbf{x}) + g(\mathbf{x}) : \ x \in \mathcal{H}\} \tag{4}$$

Assume:

(1) $f : \mathcal{H} \to \mathbb{R}$ is a smooth convex function and continuously differentiable with L_f-Lipschitz continuous gradient;

(2) $g : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is proper,convex,and lower semicontinuous;

(3) F is level bounded,$S = argminF$ is nonempty.

Main contributions

- Under certain assumptions for α_k, we show the convergence rates for the values and convergence of the iterates;
- Application to special classes of sequences α_k;

Contributions

Theorem 1.(A model example of results)

Let us make assumptions(H)

A. Supposse that the sequence $\{\alpha_k\}$ satisfies (K_0) and (K_1).

(K_0) for all $k \geq 1, \sum_{i=k}^{+\infty} \prod_{j=k}^{i} \alpha_j < +\infty$,

(K_1) for all $k \geq 1, t_{k+1}^2 - t_k^2 \leq t_{k+1}$.

Then ,sequence (x_k) generated by algorithm (IFB)

$$F(x_k) - F(x^*) = O(\frac{1}{k^2}) \quad as \;\; k \to +\infty.$$

Set $t_k = 1 + \sum_{i=k}^{+\infty} \prod_{j=k}^{i} \alpha_j$,use the convention $\prod_{j=k}^{k-1} \alpha_j = 1$,then $t_k = \sum_{i=k-1}^{+\infty} \prod_{j=k}^{i} \alpha_j$.Then t_k is well defined and $1 + \alpha_k t_{k+1} = t_k$.

Contributions

Theorem 1.

B. Assume moreover that there exists $m < 1$ such that (K_1^+)

$$t_{k+1}^2 - t_k^2 \leq m t_{k+1}$$

for every $k \geq 1$ Then,as $k \to +\infty$

$$F(x_k) - F(x^*) = o(\frac{1}{\sum_{i=1}^{k} t_i}) \ and \ \|x_k - x_{k-1}\| = o(\frac{1}{\sum_{i=1}^{k} t_i})^{\frac{1}{2}}$$

As a consequence ,we have

$$F(x_k) - F(x^*) = o(\frac{1}{t_k^2}) \ and \ \|x_k - x_{k-1}\| = o(\frac{1}{t_k}) \ as \ k \to +\infty$$

Finally, if $\alpha_k \in [0, 1]$ for every $k \geq 1$,then sequence $\{x_k\}$ converges weakly toward $\bar{x} \in argmin(F)$

Contributions

Lemma 2.

Assume hypothesis (H).
(i) For all $x, y \in \mathcal{H}$,we have[1]

$$F(y - sG_s(y)) \leq F(x) + \langle G_s(y), y - x \rangle - \frac{s}{2} \| G_s(y) \|^2$$

(ii) The operator $\mathcal{T}_s$ is nonexpansive ,the operator G_s is monotone,and the following equivalences hold true:

$$z \in argminF \Leftrightarrow z = \mathcal{T}_s(z) \Leftrightarrow G_s(z) = 0$$

Energy:

$$W_k := F(x_k) - F(x^*) + \frac{1}{2s} \| x_k - x_{k-1} \|^2$$

Where $\mathcal{T}_s = prox_{sg}(y - s\nabla f(y)), G_s(y) = \frac{1}{s}(y - \mathcal{T}_s(y))$

[1] Beck A, Teboulle M. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems[J]. Siam Journal on Imaging Sciences, 2009, 2(1): 183-202.

Contributions

Proposition 3.

Under hypothesis (H),let $\{x_k\}$ be a sequence generated by algorithm(IFB).The energy sequence$\{W_k\}$ satisfies for every $k \geq 1$,

$$W_{k+1} - W_k \leq -\frac{1-\alpha_k^2}{2s}\|x_k - x_{k-1}\|^2$$

As a consequence,the sequence $\{W_k\}$ is nonincreasing if $\alpha_k \in [0,1]$ for every $k \geq 1$.

Proof: Mainly by Lemma 2.(i)

Contributions

Proposition 4.

Under hypothesis (H)

$$h_{k+1} - h_k - \alpha_k(h_k - h_{k-1}) = \frac{1}{2}(\alpha_k^2 + \alpha_k)\|x_k - x_{k-1}\|^2 - s\langle G_s(y_k), y_k - x^*\rangle + \frac{s^2}{2}\|G_s(y_k)\|^2$$

If moreover $x^* \in argminF$, then

$$h_{k+1} - h_k - \alpha_k(h_k - h_{k-1}) \leq \frac{1}{2}(\alpha_k^2 + \alpha_k)\|x_k - x_{k-1}\|^2 - s(F(x_{k+1} - F(x^*)))$$

Where $x^* \in \mathcal{H}$,define sequence $\{h_k\}$ by $h_k = \frac{1}{2}\|x_k - x^*\|^2$.

Proof:(1) By $\|y_k - x^*\|^2$;

(2) By Lemma 2(i)

Contributions

Define $\varepsilon_k^{[1]}$

Given $x^* \in argminF$, define sequence$\{\varepsilon_k\}$ by

$$\varepsilon_k = t_k^2(F(x_k - F(x^*))) + \frac{1}{2}\|x_{k-1} + t_k(x_k - x_{k-1}) - x^*\|^2. \qquad (5)$$

Proposition 6.

Under hypothesis (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies (K_0).Let (x_k) be a sequence generated by algorithm (IFB),then

$$\varepsilon_{k+1} - \varepsilon_k \leq (t_{k+1}^2 - t_k^2 - t_{k+1})(F(x_k - F(x^*))). \qquad (6)$$

Under the assumption (K_1) $\quad t_{k+1}^2 - t_k^2 \leq t_{k+1}$ for every $k \geq 1$,then the seuence $\{\varepsilon_k\}$ is nonincreasing.

[1] Beck A, Teboulle M. A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems[J]. Siam Journal on Imaging Sciences, 2009, 2(1): 183-202.

Contributions

Proposition 7.

Under hypothesis (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies $(K_0)-(K_1)$.Let (x_k) be a sequence generated by algorithm (IFB),then

(i) For every $k \geq 1, F(x_k) - F(x^*) \leq \frac{C}{t_k^2}$ with

$$C = t_1^2(F(x_1 - F(x^*))) + \frac{1}{s}(d(x_0, S)^2 + t_1^2\|x_1 - x_0\|^2). \tag{7}$$

(ii) Assume moreover that there exists $m < 1$ such that (K_1^+) $t_{k+1}^2 - t_k^2 \leq m t_{k+1}$ for every $k \geq 1$.Then we have

$$\textstyle\sum_{k=1}^{+\infty} t_{k+1}(F(x_k) - F(x^*)) < +\infty. \tag{8}$$

Contributions

Proof:

(i) From Proposition 6,the sequence ε_k is nonincreasing.It means that $\varepsilon_k < \varepsilon_1$ for every$k \geq 1$.

$$\begin{aligned} t_k^2(F(x_k) - F(x^*)) &\leq \varepsilon_k \\ &< \varepsilon_1 = t_1^2[F(x_1) - F(x^*)] + \frac{1}{2s}\|x_0 - x^* + t_1(x_1 - x_0)\|^2 \\ &= t_1^2[F(x_1) - F(x^*)] + \frac{1}{s}(\|x_0 - x^*\|^2 + t_1^2\|(x_1 - x_0)\|^2) \end{aligned} \tag{9}$$

Since x^* can be taken arbitrarily in S, we finally obtain $t_k^2(F(x_k) - F(x^*)) \leq C$ with $C = t_1^2(F(x_1) - F(x^*)) + \frac{1}{s}(d(x_0, S))^2 + t_1^2\|x_1 - x_0\|^2$

Contributions

Proof:

(ii) By summing inequality (6) from $k = 1$ to n,we fined

$$\varepsilon_{n+1} + \sum_{k=1}^{n} (t_{k+1} - t_{k+1}^2 + t_k^2)(F(x_k) - F(x^*)) \leq \varepsilon_1$$

Since $\varepsilon_{n+1} \geq 0$ and since $t_{k+1}^2 - t_k^2 \leq m t_{k+1}$,so

$$(1 - m) \sum_{k=1}^{n} t_{k+1}(F(x_k) - F(x^*)) \leq \varepsilon_1. \tag{10}$$

The expected estimate is obtained by letting n tend to infinity.

Contributions

Proposition 8.

Under (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies $(K_0)-(K^+)$.Let (x_k) be a sequence generated by algorithm (IFB),then

$$\sum_{k=1}^{+\infty} t_k\|x_k - x_{k-1}\|^2 < +\infty. \tag{11}$$

Proof:❶ By proposition 3's inequality,multiplying this inequality by t_{k+1}^2 and summing from $k=1$ to n;

❷ Since $t_{k+1}\alpha_k = t_k - 1$ and $t_{k+1}^2 - t_k^2 \le t_{k+1}$

❸ Theorem 7 (ii)

Note:$\sum_{k=1}^{+\infty} t_{k+1}\|x_k - x_{k-1}\|^2 < +\infty$,from Theorem 7 (ii), We have

$$\sum_{k=1}^{+\infty} t_{k+1} W_k < +\infty. \tag{12}$$

Contributions

Theorem 9.(main result)

Under (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies $(K_0)-(K_1^+)$,and $\alpha_k \in [0,1]$,for every $k \geq 1$.Let (x_k) be a sequence generated by algorithm (IFB),then as $k \to +\infty$

$$F(x_k) - F(x^*) = o(\frac{1}{\sum_{i=1}^{k} t_i}) \quad and \quad \|x_k - x_{k-1}\| = o(\frac{1}{\sum_{i=1}^{k} t_i})^{\frac{1}{2}}. \tag{13}$$

As a consequence ,we have

$$F(x_k) - F(x^*) = o(\frac{1}{t_k^2}) \quad and \quad \|x_k - x_{k-1}\| = o(\frac{1}{t_k}). \tag{14}$$

hence $\lim_{k\to+\infty} F(x_k) = minF$,and $\lim_{k\to+\infty} \|x_k - x_{k-1}\| = 0$.

Contributions

Proof:

❶ Sequence $\{W_k\}$ is nonincreasing because $\alpha_k \in [0,1]$ for every $k \geq 1$;
❷ $\sum_{k=1}^{+\infty} t_{k+1} W_k < +\infty \Rightarrow \sum_{k=1}^{+\infty} t_k W_k < +\infty$.
❸ By following Lemma

Lemma

Let τ_k be a nonnegative sequence such that $\sum_{k=1}^{+\infty} \tau_k = +\infty$.Assume that δ_k is a nonnegative and nonincreasing sequence satisfying $\sum_{k=1}^{+\infty} \tau_k \delta_k < +\infty$.Then

$$\delta_k = o\left(\frac{1}{\sum_{i=1}^{k} \tau_i}\right) \quad as \ k \to +\infty. \tag{15}$$

$\Rightarrow W_k = o\left(\frac{1}{\sum_{i=1}^{k} t_i}\right)$ $as \ k \to +\infty$
(2) By K_1,and summing from $i = 1 \ to \ k-1$

$$t_k^2 \leq t_1^2 + \sum_{i=1}^{k-1} t_{i+1} = t_1^2 - t_1 + \sum_{i=1}^{k} t_i$$

Contributions

Theorem 10.(main result)

Under (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies (K_0)-(K_1^+),and $\alpha_k \in [0,1]$,for every $k \geq 1$.Then any sequence (x_k) be a sequence generated by algorithm (IFB) converges weakly,and its limit belongs to $argminF$.

Proof:Main by the following Lemma

Lemma 21.

Let S be a nonempty subset of $\mathcal{H}$ and $\{x_k\}$ a sequence of elements of $\mathcal{H}$ .Assume that

(i) every sequential weak cluster point of $\{x_k\}$,as $k \rightarrow +\infty$ belongs to S;

(ii) for every $z \in S$,$\lim_{k\rightarrow+\infty} \|x_k - z\|$ exists.

Then $\{x_k\}$ converges weakly as $k \rightarrow +\infty$ to a point in S.

Contributions

Proof:Main by the following Lemma

Lemma 23.

Given a nonnegative sequence $\{\alpha_k\}$ satisfying K_0,let $t_k = 1 + \sum_{i=k}^{+\infty} \prod_{j=k}^{i} \alpha_j$.Let $\{a_k\}$ and $\{\omega_k\}$ be two sequences of nonnegative numbers such that

$$\alpha_{k+1} \leq \alpha_k a_k + \omega_k. \tag{16}$$

for all $k \geq 0$,if $\sum_{k=0}^{+\infty} t_{k+1}\omega_k < +\infty$,then $\sum_{k=0}^{+\infty} a_k < +\infty$

Contributions

Theorem 11.

Under (H),assume that the nonnegative sequence $\{\alpha_k\}$ satisfies (K_0)-(K_1^+),and $\alpha_k \in [0,1]$,for every $k \geq 1$.Suppose moreover that the function F is even.Then any sequence (x_k) be a sequence generated by algorithm (IFB) converges strongly,and its limit belongs to $argminF$.

Contributions:To special classes of sequences α_k

Proposition 14.

Let $c \in [0,1]$ and let $\alpha_k \in [0,1]$ together with inequality $\frac{1}{1-\alpha_{k+1}} - \frac{1}{1-\alpha_k} \leq c$ for every $k \geq 1$.Then the following holds true:

(i) Condition K_0 is satisfied and have for every $k \geq 1$

$$t_{k+1} \leq \frac{1}{(1-c)(1-\alpha_k)}$$

(ii) If $c \leq \frac{1}{3}(resp., c < \frac{1}{3})$,then condition $K_1(resp., (K_1^+))$ is fulfilled.

Contributions

Proposition 15.

Let α_k be a sequence such that $\alpha_k \in [0,1]$ for every $k \geq 1$.Given $c \in [0,1]$ assume that

$$\lim_{k\to+\infty} \frac{1}{1-\alpha_{k+1}} - \frac{1}{1-\alpha_k} = c. \tag{17}$$

Then

$$t_{k+1} \sim \frac{1}{(1-c)(1-\alpha_k)} \;\; as \;\; k \to +\infty. \tag{18}$$

Contributions

Theorem 16.(main result)

Under hypothesis (H),assume that the sequence α_k is such that $\alpha_k \in [0,1]$ for every $k \geq 1$.Let x_k be a sequence generated by (IFB). Then we have following:

(i) $F(x_k) - F(x^*) = O(\frac{1}{t_k^2})$ as $k \to +\infty$; by Theorem 7 (i)

(ii) $\sum_{k=1}^{+\infty} t_{k+1}(F(x_k) - F(x^*)) < +\infty$; by Theorem 7 (ii)

(iii) $\sum_{k=1}^{+\infty} t_k \|x_k - x_{k-1}\|^2 < +\infty$; by Proposition 8

(iv) $F(x_k) - F(x^*) = o(\frac{1}{\sum_{i=1}^{k} t_i})$ and $\|x_k - x_{k-1}\| = o(\frac{1}{\sum_{i=1}^{k} t_i})^{\frac{1}{2}}$ as $k \to +\infty$; by Theorem 9

(v) sequence x_k converges weakly as $k \to +\infty$ toward $argminF$ .If the F is even ,then the convergence is strong in $\mathcal{H}$; by Theorem 10 and Theorem 11

Contributions

Corollary 17.

Under hypothesis (H).Given $\alpha \geq 3$,suppose that $\alpha_k = 1 - \frac{\alpha}{k}$ for every $k \geq 1$.Let x_k be a sequence generated by (IFB). Then we have following:

(i) $F(x_k) - F(x^*) = O(\frac{1}{k^2})$ as $k \to +\infty$;

Assuming now $\alpha > 3$:

(ii) $\sum_{k=1}^{+\infty} k(F(x_k) - F(x^*)) < +\infty$;

(iii) $\sum_{k=1}^{+\infty} k\|x_k - x_{k-1}\|^2 < +\infty$;

(iv) $F(x_k) - F(x^*) = o(\frac{1}{k^2})\ and\ \|x_k - x_{k-1}\| = o(\frac{1}{k})\ as\ k \to +\infty$;

(v) sequence x_k converges weakly as $k \to +\infty$ toward $argminF$ .If the F is even ,then the convergence is strong in $\mathcal{H}$;

Proof: Mainly by Theorem 16.

Contributions

Corollary 18.

Under hypothesis (H).Given $\theta > 0$,suppose that $\alpha_k = 1 - \frac{(\ln k)^\theta}{k}$ for every $k \geq 1$.Let x_k be a sequence generated by (IFB). Then we have following:

(i) $\sum_{k=1}^{+\infty} \frac{k}{(\ln k)^\theta}(F(x_k) - F(x^*)) < +\infty$;

(ii) $\sum_{k=1}^{+\infty} \frac{k}{(\ln k)^\theta}\|x_k - x_{k-1}\|^2 < +\infty$;

(iii) $F(x_k) - F(x^*) = o(\frac{(\ln k)^\theta}{k^2})$ *and* $\|x_k - x_{k-1}\| = o(\frac{(\ln k)^\theta}{k^2})^{1/2}$ *as* $k \to +\infty$;

(iv) sequence x_k converges weakly as $k \to +\infty$ toward $argminF$ .If the F is even ,then the convergence is strong in $\mathcal{H}$;

Proof: Mainly by Theorem 16.

Contributions

Corollary 19.

Under hypothesis (H).Given $\alpha \geq 0$ and $r \in [0,1]$,suppose that $\alpha_k = 1 - \frac{\alpha}{k^r}$ for every $k \geq 1$.Let x_k be a sequence generated by (IFB). Then we have following:

(i) $\sum_{k=1}^{+\infty} k^r (F(x_k) - F(x^*)) < +\infty$;

(ii) $\sum_{k=1}^{+\infty} k^r \|x_k - x_{k-1}\|^2 < +\infty$;

(iii) $F(x_k) - F(x^*) = o(\frac{1}{k^{1+r}})$ and $\|x_k - x_{k-1}\| = o(\frac{1}{k^{1+r}})^{1/2}$ as $k \to +\infty$;

(iv) sequence x_k converges weakly as $k \to +\infty$ toward $argminF$ .If the F is even ,then the convergence is strong in $\mathcal{H}$;

Proof: Mainly by Theorem 16.

Contributions

Corollary 20.

Under hypothesis (H).Given $\alpha_k \in [m, M]$ for every $k \geq 1$ with $0 \leq m \leq M$,suppose that $\frac{1-m}{1-M} \leq \frac{3}{2}$ .Let x_k be a sequence generated by (IFB). Then we have following: (i) $\sum_{k=1}^{+\infty}(F(x_k) - F(x^*)) < +\infty$;

(ii) $\sum_{k=1}^{+\infty} \|x_k - x_{k-1}\|^2 < +\infty$;

(iii) $F(x_k) - F(x^*) = o(\frac{1}{k})$ *and* $\|x_k - x_{k-1}\| = o(\frac{1}{k^{1/2}})$ *as* $k \to +\infty$;

(iv) sequence x_k converges weakly as $k \to +\infty$ toward $argminF$ .If the F is even ,then the convergence is strong in $\mathcal{H}$;

Proof: Mainly by Theorem 16.

Summary

α_k	$\alpha_k = 1 - \frac{\alpha}{k},\ \alpha > 3$	$\alpha_k = 1 - \frac{3}{k}$	$\alpha_k = 1 - \frac{(\ln k)^\theta}{k}$	$\alpha_k = 1 - \frac{\alpha}{k^r},\ r \in]0,1[$	$\alpha_k \to l \in [0,1[$
t_k	$\frac{k-1}{\alpha-1}$	$\frac{k-1}{2}$	$\sim \frac{k}{(\ln k)^\theta}$	$\sim \frac{k^r}{\alpha}$	$\sim \frac{1}{1-l}$
W_k	$o\left(\frac{1}{k^2}\right)$	$\mathcal{O}\left(\frac{1}{k^2}\right)$	$o\left(\frac{(\ln k)^\theta}{k^2}\right)$	$o\left(\frac{1}{k^{r+1}}\right)$	$o\left(\frac{1}{k}\right)$

Figure 1

❶ $\alpha_k = 1 - \frac{\alpha}{k}, \alpha > 3 \Leftrightarrow \alpha_k = \frac{k-1}{k+\alpha-1}, \alpha > 3$;

❷ $\alpha_k = 1 - \frac{3}{k} \Leftrightarrow t_1 = 1, t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}, \alpha_k = \frac{t_k - 1}{t_{k+1}}$. The convergence of the iterative sequence is still not given.

2

- Wen B, Xue X. On the convergence of the iterates of proximal gradient algorithm with extrapolation for convex nonsmooth minimization problems[J]. Journal of Global Optimization, 2019, 75(3): 767-787.

Background

we consider the following problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n}\{F(x) := f(\mathbf{x}) + g(\mathbf{x})\} \tag{19}$$

Assume:

(1) f is a smooth convex function and continuously differentiable with L_f-Lipschitz continuous gradient.

(2) g be a proper closed convex function.

(3) F is level bounded .

Main contributions

- First show that for a large class of extrapolation parameters including the extrapolation parameters chosen in FISTA the successive changes of iterates go to 0;
- Second, based on the Lojasiewicz inequality, we establish the global convergence of iterates generated by the proximal gradient algorithm with extrapolation with an additional assumption on the extrapolation coefficients;
- Third, proved the length of the iterates is finite.

Contributions

Define auxiliary sequence

$$H_k = F(x_k) + \frac{L}{2}\|x_k - x_{k-1}\|^2 \tag{20}$$

Lemma 1

Let x_k be a sequence generated by PG_e. Then the following statements hold.

(i) The sequence $\{H_k\}$ is nonincreasing.

(ii) The sequence $\{x_k\}$ is bounded.

(iii) The sequence $\{H_k\}$ is convergent.

(iv)$\sum_{k=0}^{\infty}(\frac{L}{2} - \frac{L}{2}\beta_{k+1}^2)\|x_{k+1} - x_k\|^2 < \infty$.

Contributions

Proof: (i): Mainly by ❶ $G(x) = \langle \nabla f(y_k), x\rangle + \frac{L}{2}\|x - y_k\|^2 + g(x)$ is strong convex; ❷ ∇f is Lipschitz continuous and f is convex.

(ii) :$F(x_k) \leq H_k \leq H_0 < \infty$,$F$ is level bounded.

(iii) :H_k is bounded from below and nonincreasing.

(iv) :$H_{k+1} - H_k \leq (\frac{L}{2}\beta_k^2 - \frac{L}{2})\|x_k - x_{k-1}\|^2$,summing both sides from 1 to N,and letting $N \to \infty$,and H_k is convergent.Then the infinite sum exists and is finite.

Contributions

Lemma 2

Let $\alpha_k \in (0,1)$ be a nondecreasing nonnegative sequence with $\lim_{k\to+\infty} \alpha_k = 1$

Then

$$\sum_{k=1}^{\infty} |\alpha_{k+1}(1-\alpha_{k+2}) - (1+\alpha_k)(1-\alpha_{k+1}) + 1 - \alpha_k| < \infty \tag{21}$$

Lemma 3

Suppose that the sequence $\{\beta_k\}$ in PG_e is nondecreasing with $\sum_{k=1}^{\infty}(1-\beta_k) = \infty$.Let $\{x_k\}$ be a sequence generated by PG_e.Then $\lim_{k\to\infty} \|x_k - x_{k-1}\| = 0$

Contributions

Proof:

❶ for $sup_k \beta_k < 1$ by Lemma 1.(iv) is immediately;

❷ for $sup_k \beta_k = 1$, equivalent to prove $\lim_{k\to\infty} H_k = F(x^*)$ by auxiliary sequence $h_k = \|x_k - z\|^2, z \in \mathcal{X}$ and Lemma 1,$\mathcal{X}$ denote the set of stationary points of F.

Contributions

Corollary 1

Let $\{x_k\}$ be a sequence generated by FISTA for solving problem. Then $\lim_{k\to\infty} \|x_{k+1} - x_k\| = 0$.

Corollary 2

Suppose that the sequence $\{\beta_k\}$ in PG_e is nondecreasing with $\sum_{k=1}^{\infty}(1-\beta_k) = \infty$. Let $\{x_k\}$ be sequence generated by PG_e. Then $\lim_{k\to\infty} F(x_k) = min\ F$ and $\lim_{k\to\infty} k(1-\beta_k)(H_k - minF) = 0$

Lemma 4

Suppose that the sequence $\{\beta_k\}$ in PG_e is nondecreasing with $\sum_{k=1}^{\infty}(1-\beta_k) = \infty$. Let $\{x_k\}$ be sequence generated by PG_e. Then any accumulation point of $\{x_k\}$ is a minimizer of F.

Contributions

Definition 1(Lojasiewicz property)

A function F is said to satisfy the Lojasiewicz property at $\hat{x} \in dom\partial F$ if there exist $c > 0, \epsilon > 0.a > 0$ and $\theta \in [0, 1)$, such that

$$c \cdot dist(0, \partial(F(x))) \geq (F(x) - F(\hat{x}))^{\theta} \tag{22}$$

where $\|x - \hat{x}\| \leq \epsilon$,$F(\hat{x}) < F(x) < F(\hat{x}) + a$.Then we say that the Lojasiewicz exponent of F at $\hat{x}$ is θ.If F satisfies above inequality at all points in dom ∂F with θ,then we say that F satisfies the uniformized Lojasiewicz property.

Contributions

Theorem 1

Suppose that F satisfies the uniformized Lojasiewicz property with the Lojasiewicz exponent $\frac{1}{2} \leq \theta < 1$,the sequence $\{\beta_k\}$ in PG_e is nondecreasing with $\lim_{k\to\infty} \beta_k = 1$ and $\lim_{k\to\infty} k(1-\beta_k)^{1+\frac{2}{1-\theta}} \geq 1$.Let $\{x_k\}$ be a sequence generated by PG_e.Then $\{x_k\}$ converges to a minimizer of F.Moreover,the length of $\{x_k\}$ is finite.

3

- Liu H , Wang T , Liu Z . Convergence Rate of Inertial Forward-Backward Algorithms Based on the Local Error Bound Condition[J]. 2020.

Background

we consider the following problem:

$$\min_{\mathbf{x}\in H}\{F(x) := f(\mathbf{x}) + g(\mathbf{x})\} \tag{23}$$

Assume:

(1)H be a real Hilbert space.

(2)$f : H \rightarrow R$ be a smooth convex function and continuously differentiable with L_f-Lipschitz continuous gradient.

(3)$g : H \rightarrow R \cup +\infty$ be a proper lower semi-continuous convex.

(4)$X^* := argmin\ F \neq \emptyset$,$x^* \in X^*$, $F^* := F(x^*)$.

Background

Algorithm 1 (Inertial Forward-Backward algorithm (IFB))

Step 0: Take $y_1 = x_0 \in R^n$, $t_1 = 1$.Input $\lambda = \frac{\mu}{L_f}$,where $\mu \in [0, 1]$

Step k: Commpute:

$$x_k = p_{\lambda g}(y_k - \lambda \nabla f(y_k))$$

$y_{k+1} = x_k + \gamma_k(x_k - x_{k-1})$ where $\gamma_k = \frac{t_k - 1}{t_{k+1}}$.

Contributions

Based on the following inequality[2]:

$$F(p_{\lambda g}(y)) - F(x) \leq \frac{1}{2\lambda}\|x-y\|^2 - \frac{1}{2\lambda}\|p_{\lambda g}(y)-x\|^2. \tag{24}$$

Lemma 2.1.

For any $y \in R^n$,$\lambda = \frac{\mu}{L_f}$,where $\mu \in [0,1]$

$$F(p_{\lambda g}(y)) \leq F(x) + \frac{1}{2\lambda}\|x-y\|^2 - \frac{1-\mu}{2\lambda}\|p_{\lambda g}(y)-y\|^2 - \frac{1}{2\lambda}\|p_{\lambda g}(y)-x\|^2. \tag{25}$$

[2] Chambolle, A. & Dossal, Ch. (2015). On the Convergence of the Iterates of the "Fast Iterative Shrinkage/Thresholding Algorithm". Journal of Optimization Theory and Applications. 166.

Contributions

Assumption A_0.

For any $\xi_0 \geq F^*$,there exist a $\varepsilon_0 > 0$ and $\tau_0 > 0$ such that

$$dist(x, X^*) \leq \tau_0. \tag{26}$$

whenever $\|p_{\frac{1}{L_f}g}(x) - x\| < \varepsilon_0$ and $F(x) \leq \xi_0$.

Lemma 2.2

Boundedness of level sets[4],for $\lambda_1 \geq \lambda_2 > 0$

$$\|p_{\lambda_1 g}(x) - x\| \geq \|p_{\lambda_2 g}(x) - x\| \quad and \quad \frac{\|p_{\lambda_1 g}(x) - x\|}{\lambda_1} \leq \frac{\|p_{\lambda_2 g}(x) - x\|}{\lambda_2} \tag{27}$$

[4] Bnouhachem A. A self-adaptive method for solving general mixed variational inequalities[J]. Journal of Mathematical Analysis and Applications, 2005, 309(1): 136-150.

Contributions (Main result)

Theorem 2.1

Let x_k,y_k be generated by Algorithm 1.Supose that exists a positive interger k_0 such that for $k > k_0$,$0 \leq \gamma_k \leq 1$ .Then,

(1)$\sum_{k=1}^{\infty} \|x_{k+1} - y_{k+1}\|^2$ is convergent.

(2)$\lim_{k\to\infty} F(x_k) = F(x^*)$.

Proof: (1) Main by Lemma 2.1;

(2) Main by Lemma 2.2,nonexpansiveness property of the proximal operator and ∇f is Lipschitz continuous,$\{F(x_k) + \frac{1}{2\lambda}\|x_{k+1} - x_k\|^2\}$ is nonincreasing ,Assumption A_0 and result 1 .

Contributions

Assumption A_1

(Local error bound condition[5]) For any $\xi \geq F^*$,there exist a $\varepsilon > 0$ and $\bar{\tau}$>0 such that

$$dist(x, X^*) \leq \bar{\tau} \| p_{\frac{1}{L_f} g}(x) - x \|. \tag{28}$$

whenever $\| p_{\frac{1}{L_f} g}(x) - x \| < \varepsilon$ and $F(x) \leq \xi$

Assumption A_2

$$\lim_{k \to \infty} t_k = +\infty$$

Assumption A_3

There exists a positive constant p such that$\lim_{k \to \infty} k^p (\frac{t_{k+1}}{t_k} - 1) = c$,where $c > 0$.

[5] Wen B, Chen X, Pong T K, et al. Linear Convergence of Proximal Gradient Algorithm with Extrapolation for a Class of Nonconvex Nonsmooth Minimization Problems[J]. Siam Journal on Optimization, 2017, 27(1): 124-145.

Contributions

Lemma 2.3

Suppose that Assumptions $A_1 - A_3$ hold.Let x_k be generated by Algorithm 1 and $x^* \in X^*$.There exists a constant $\tau_1 > 0$ such that

$$\forall k \geq 1, \quad F(x_{k+1}) - F(x^*) \leq \frac{\tau_1}{\lambda} \|y_{k+1} - x_{k+1}\|^2. \tag{29}$$

Proof: Similar with the proof of Theorem 2.1.

Lemma 2.4(main)

Suppose that Assumptions$A_1 - A_3$ hold .Let x_k be generated by Algorithm 1 and $x^* \in X^*$.Then,$\sum_{k=1}^{\infty} t_{k+1}^2 (F(x_{k+1}) - F(x^*))$ is convergent and $\|x_k - x_{k-1}\| \leq O(\frac{1}{t_k})$.

Note: Lemma 2.4 implies that $F(x_k) - F(x^*) = o(\frac{1}{t_k^2})$

Contributions

Proof:

❶ Applying Lemma 2.1 and let $x := (1 - \frac{1}{t_{k+1}})x_k + \frac{1}{t_{k+1}}x^*, y := y_{k+1}$.

❷ Deonte that $\phi_k = (t_k^2 + \rho_k)(F(x_k) - F(x^*)) + \frac{1}{2\lambda}\|\mu_k\|$,where $\mu_k = t_k x_k - (t_k - 1)x_{k-1} - x^*$,$\rho_k = t_{k+1}^2 - t_k^2 - t_{k+1}$,

$$\phi_k - \phi_{k+1} \geq \frac{1-\mu}{4\lambda} t_{k+1}^2 \|x_{k+1} - y_{k+1}\|^2$$

show that the ϕ_k is nonincreasing and bound below,so ϕ_k is convergent ;

❸ By Lemma 2.3 Then,$\sum_{k=1}^{\infty} t_{k+1}^2 (F(x_{k+1}) - F(x^*))$ is convergent.

For $\|x_k - x_{k-1}\| \leq O(\frac{1}{t_k})$

❶ by $\frac{\rho_k}{t_k^2} = 0$,have $\lim_{k\to\infty}(t_k^2 + \rho_k)(F(x_k) - F(x^*)) = 0$,because $\{\phi_k\}$ is convergent,then μ_k is convergent;

❷ i.e. $\|x_k - x^* - (1 - \frac{1}{t_k})(x_{k-1} - x^*)\| \leq \frac{\epsilon}{t_k}$,then $\|x_k - x^*\| \leq O(\frac{1}{t_k})$

Contributions

Theorem 2.2 (Main result)

Suppose that Assumptions $A_1 - A_3$ hold and $\sum_{k=1}^{\infty} \frac{1}{t_k}$ is convergent. Then,the iterates x_k strongly converges to a minimizer of F

Proof:In the proof of Lemma 2.4,have $\exists \epsilon > 0$ such that $\|x_k - x_{k-1}\| \leq \frac{\epsilon}{t_k}$,because $\sum_{k=1}^{\infty} \frac{1}{t_k}$ is convergent,then $\{x_k\}$ is a Cauchy series and since F is lsc and convex .

In the following, we consider the case that $\sum_{k=1}^{\infty} \frac{1}{t_k}$ is divergent:

Assumption A^+

There exist $0 < M < 2$ and $m > 0$ such that $t_{k+1} - t_k \leq M, \forall k > m$

Note: $t_{k+1} \leq t_k + M, \sum_{k=1}^{\infty} \frac{1}{t_{k+1}} \geq \sum_{k=1}^{\infty} \frac{1}{t_1 + kM}$, where $\sum_{k=1}^{\infty} \frac{1}{t_1 + kM}$ is divergent,so $\sum_{k=1}^{\infty} \frac{1}{t_k}$ also divergent.

Contributions

Lemma 2.5 (Main result)

Suppose that Assumptions $A_1 - A_3$ and A^+ hold .Let x_k be generated by Algorithm 1 and $x^* \in X^*$.Then,$(F(x_{k+1}) - F(x^*)) = o(\frac{1}{t_k^3})$ and $\|x_k - x_{k-1}\| \leq o(\frac{1}{t_k^{1.5}})$

Proof: Main by denote $\psi_k = F(x_k) - F(x^*) + \frac{1}{2\lambda}\|x_k - x_{k-1}\|^2$,prove that $\lim_{k\to\infty} t_k^3\psi_k = 0$ by sequence $\{t_k^3\psi_k\}$ and Lemma 2.4.

Lemma 2.6 (for the sequence $\{x_k\}$ weakly converges)

Let X^* be a nonempty subset of H and x_k a sequence of elements of $\mathcal{H}$ .Assume that

(1) every sequential weak cluster point of x_k,as $k \to \infty$,belongs to X^*;

(2) for every $x^* \in X^*$,$\lim_{k\to\infty} \|x_k - x^*\|$ exists. Then x_k converges weakly to point in X^* as $k \to \infty$.

Contributions

Theorem 2.3

Suppose that Assumptions $A_1 - A_3$ and A^+ hold.We have that the sequence x_k generated by Algorithm 1 converges weakly to its limit belongs to X^*.

Theorem 2.4

Suppose that Assumptions $A_1 - A_3$ and A^+ hold and $\sum_{k=1}^{\infty} \frac{1}{t_k^{1.5}}$ is convergent.Then, the iterates x_k strongly converges to a minimizer of F.

Proof:In the proof of Lemma 2.5,$\lim_{k\to\infty} t_k^{1.5}\|x_k - x_{k-1}\| = 0$,i.e.,$\exists\epsilon > 0$,

$$\|x_k - x_{k-1}\| \leq \frac{\epsilon}{t_k^{1.5}}$$

because $\sum_{k=1}^{\infty} \frac{1}{t_k^{1.5}}$ is convergent,$\{x_k\}$ is a Cauchy series ,combining with Theorem 2.3, $\{x_k\}$ strongly converges to $\bar{x} \in X^*$

Contributions: several options for t_k

Corollary 3.1

Suppose that Assumption A_1 holds, $t_1 = 1$ and $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$. Let x_k be generated by Algorithm 1 and $x^* \in X^*$. Then:

(1) $F(x_k) - F(x^*) = o(\frac{1}{k^3})$ and $\|x_k - x_{k-1}\| = o(\frac{1}{k^{1.5}})$.

(2) x_k is converges sublinearly to $\bar{x} \in X^*$ at the $o(\frac{1}{k^{0.5}})$ rate of convergence.

Proof: Mainly by Lemma 2.5 and Theorem 2.4.

Contributions: several options for t_k

Corollary 3.2

Suppose that Assumption A_1 holds, $t_k = \frac{k^r - 1 + \alpha}{\alpha}$ where $0 < r \leq 1$ and $\begin{cases} a \geq 1, \text{ if } r = 1 \\ a > 0, \text{ if } r \neq 1 \end{cases}$. Suppose that Assumption A_1 holds.Let x_k be generated by Algorithm 1 with $t_k = \frac{k^r - 1 + \alpha}{\alpha}$ $(0 < r \leq 1)$ and $x^* \in X^*$,Then

(1) For any positive integer p, $F(x_k) - F(x^*) = o(\frac{1}{k^{p(1-r)+2r+1}})$ and $\|x_k - x_{k-1}\| = o(\frac{1}{k^{0.5p(1-r)+r+0.5}})$;

(2) For any positive integer p, x_k is converges sublinearly to $\bar{x} \in X^*$ at the $o(\frac{1}{k^{0.5p(1-r)+r-0.5}})$ rate of convergence.

Proof: Mainly by Lemma 2.4, $A_1 - A_3$ hold.

Contributions: several options for t_k

Corollary 3.3

Suppose that Assumption A_1 holds, $t_k = \frac{k^r - 1 + \alpha}{\alpha} (r > 1)$, and $x^* \in X^*$. Then:

(1) $F(x_k) - F(x^*) = o(\frac{1}{k^{2r}})$ and $\|x_k - x_{k-1}\| = O(\frac{1}{k^r})$.

(2) x_k is converges sublinearly to $\bar{x} \in X^*$ at the $O(\frac{1}{k^{r-1}})$ rate of convergence.

Proof: (main idea) Assumptions $A_1 - A_3$ hold. $\lim_{k \to \infty} \frac{t_k}{k^r} = \frac{1}{\alpha}$ and $\sum_{k=1}^{\infty} \frac{1}{t_k}$ is convergent, we can deduce that the result 1) holds by Lemma 2.4 and x_k strongly converges to $\bar{x} \in X^*$ by Theorem 2.2.

Contributions: several options for t_k

Corollary 3.4

Suppose that Assumption A_1 holds, $t_k = \frac{k}{\ln^{\theta} k} (k > 1)$, where $t_1 = 1$ and $\theta > 0$.

Let x_k be generated by Algorithm 1 and $x^* \in X^*$. Then :

(1) $F(x_k) - F(x^*) = o(\frac{\ln^{3\theta} k}{k^3})$ and $\|x_k - x_{k-1}\| = o(\frac{\ln^{1.5\theta} k}{k^{1.5}})$.

(2) x_k is converges sublinearly to $\bar{x} \in X^*$ at the $o(\frac{\ln^{1.5\theta} k}{k^{0.5}})$ rate of convergence.

Proof: (main idea) by Lemma 2.5 and Theorem 2.4.

Contributions: several options for t_k

Corollary 3.5

Suppose that Assumption A_1 holds, $t_k = e^{(k-1)^\alpha} 0 < \alpha < 1$. Let x_k be generated by Algorithm 1 and $x^* \in X^*$. Then :

(1) $F(x_k) - F(x^*) = o(\frac{1}{e^{2(k-1)^\alpha}})$ and $\|x_k - x_{k-1}\| = O(\frac{1}{k^{e^{(k-1)^\alpha}}})$.

(2) x_k is converges sublinearly to $\bar{x} \in X^*$ at the $O((k-1)^{\alpha\lceil \frac{1}{\alpha} - 1\rceil} e^{-(k-1)^\alpha})$ rate of convergence.

Proof: (main idea) by Lemma 2.4 .

Summary

- Based on the local error bound condition,exploit some assumption conditions for the important parameter t_k in IFB;
- Discuss the convergence results including convergence rate of function value and strong or weak convergence of iterates generated by the corresponding IFB;
- Discuss four choices of t_k , which includes the ones in original FISTA.

Numerical Experiments

For Corollary 3.3,$t_k = \frac{k^r - 1 + \alpha}{\alpha}$.the rate of convergence should improve constantly as r increasing,set $\alpha = 4, r = 2, r = 4, r = 6, r = 8$,and consider LASSO problem

$$\min_{x \in \mathcal{R}^n} F(x = \frac{1}{2}\|Ax - b\|^2 + \delta\|x\|_1$$

,where $A \in \mathcal{R}^{m \times n}$ be a Gaussian matrix ,set $\delta = 1$

Table 1: Different r for $t_k = \frac{k^r - 1 + \alpha}{\alpha}$

(m,n)	r=2(iter,time)	r=4(iter,time)	r=6(iter,time)	r=8(iter,time)
(100,1000)	(67,0.1541)	(46,0.0885)	(40,0.0637)	(24,0.0392)
(200,2000)	(91,0.6288)	(61,0.3874)	(47,0.2685)	(38,0.1846)
(500,5000)	(107,4.0735)	(74,2.8082)	(38,1.3771)	(32,1.1198)

Numerical Experiments

Consider the following five algorithms;

(1) FISTA;

(2) FISTA_CD with $\alpha = 4$;

(3) FISTA_(8),i.e.,$t_k = \frac{k^r - 1 + \alpha}{\alpha}$ $(r = 8 \ and \ \alpha = 4)$;

(4) FISTA_(0.5),i.e.,$t_k = \frac{k^r - 1 + \alpha}{\alpha}$ $(r = 0.5 \ and \ \alpha = 4)$;

(5) FISTA_exp,i.e.,$t_k = e^{(k-1)^\alpha} 0 < \alpha < 1$,and set $\alpha = 0.5$.

Table 2: Different t_k for IFB

(m,n)	FISTA	FISTA_CD	FISTA_(8)	FISTA_(0.5)	FISTA_exp
(100,1000)	(90,0.2149)	(50,0.0837)	(38,0.0762)	(53,0.0841)	(48,0.0765)
(200,2000)	(88,0.6709)	(41,0.3583)	(38,0.2464)	(57,0.3512)	(58,0.3359)
(500,5000)	(113,3.9531)	(47,1.6478)	(40,1.5632)	(62,2.4025)	(72,2.5331)

Thank you!