Research Article

Convergence Analysis of a Modified Forward-Backward Splitting Algorithm for Minimization and Application to Image Recovery

Kunrada Kankam, Watcharaporn Cholamjiak, and Prasit Cholamjiak

School of Science, University of Phayao, Phayao 56000, Thailand

Correspondence should be addressed to Prasit Cholamjiak; prasitch2008@yahoo.com

Received 4 August 2022; Revised 16 August 2022; Accepted 8 September 2022; Published 6 October 2022

Academic Editor: Naeem Jan

Copyright © 2022 Kunrada Kankam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many applications in applied sciences and engineering can be considered as the convex minimization problem with the sum of two functions. One of the most popular techniques to solve this problem is the forward-backward algorithm. In this work, we aim to present a new version of splitting algorithms by adapting with Tseng’s extragradient method and using the linesearch technique with inertial conditions. We obtain its convergence result under mild assumptions. Moreover, as applications, we provide numerical experiments to solve image recovery problem. We also compare our algorithm and demonstrate the efficiency to some known algorithms.

1. Introduction and Preliminaries

In various fields of applied sciences and engineering such as signal recovery, image restoration and machine learning [1–9] can be formulated as convex minimization problem (CMP) in the term of sum of nonsmooth and smooth functions. Let $H$ be a real Hilbert space. CMP is modeled as follows:

$$\min \{f(k) + g(k): k \in H\},$$  \hspace{1cm} (1)

where $f : H \rightarrow (-\infty, +\infty]$ and $g : H \rightarrow \mathbb{R}$ are two proper lower semicontinuous convex functions such that $f$ is differentiable on $H$. For any $\lambda > 0$, it is known that $k_*$ is an optimal solution to (1) if

$$0 \in \lambda \nabla f(k_*) + \lambda \partial g(k_*) \iff (I - \lambda \nabla f)(k_*) \in (I + \lambda \partial g)(k_*) \iff k_* = (I + \lambda \partial g)^{-1}(k_* - \lambda \nabla f(k_*)),$$ \hspace{1cm} (2)

where $\nabla f$ is the gradient of $f$ is linear function, which is defined by

$$\langle \nabla f(k), y \rangle = f'(k, y),$$ \hspace{1cm} (3)

where the derivative of $f$ at $k$ in the direction $y$ is $f'(k, y) = (f(k + ty) - f(k))/t$ and $\partial g(\cdot)$ is the classical subdifferential of $g$ which is given by

$$\partial g(\cdot) = \{z \in H : g(u) - g(\cdot) - \langle z, u - (\cdot) \rangle \geq 0, \forall u \in \mathbb{R}^n\}.$$ \hspace{1cm} (4)

It is known that $\partial g$ is maximal monotone and if $g$ is differentiable, then $\partial g$ is the gradient of $g$ denoted by $\nabla g$. This leads to the classical forward-backward splitting algorithm (FBS) [10, 11] which is defined by $k^n \in H$ and

$$k^{n+1} = \text{prox}_{\lambda_n g}(k^n - \lambda_n \nabla f(k^n)),$$ \hspace{1cm} (5)

where $\lambda_n > 0$ and $\text{prox}_{\lambda_n g} = (I + \lambda_n \partial g)^{-1}$ is the proximal operator. On the one hand, (5) includes the gradient algorithm $k^{n+1} = k^n - \lambda_n \nabla f(k^n)$, where $\lambda_n > 0$ and $f$ is a Lipschitz continuous gradient. Moreover, (5) includes the proximal point algorithm $k^{n+1} = \text{prox}_{\lambda_n g} k^n$, where $\lambda_n > 0$ and $g$ is a nondifferentiable function. We know that the proximal operator is single-valued and is characterized by

$$k - \frac{\text{prox}_{\lambda_n g}(k)}{\lambda} \in \partial g(\text{prox}_{\lambda_n g}(k)),$$ \hspace{1cm} (6)
for all \( k \in H \) and \( \lambda > 0 \). The iteration (5) has been attracted extensively by many researchers. See, for example, [12–18]. One popular method for solving (1) is the modified forward-backward splitting method (MFBS) or Tseng’s extragradient method [19]; MFBS is generated by \( k^0 \in H \) and

\[
k^{n+1} = \text{prox}_{\alpha_n}(k^n - \alpha_n \nabla f(k^n)) - \alpha_n (\nabla f(\text{prox}_{\alpha_n}(k^n - \alpha_n \nabla f(k^n))) - \nabla f(k^n)).
\]

(7)

where \((\alpha_n) \subset (0, +\infty)\) is a real sequence. The convergence rate is well known for the speed of \( O(1/n) \). Later, various schemes were proposed to improve the convergence and accelerate the method. Among them, Lorenz and Pock [20] have improved the convergence speed of FBS from the standard \( O(1/n) \) to \( O(1/n^2) \).

Recently, Beck and Teboulle [21] introduced a fast iterative shrinkage-thresholding algorithm (FISTA-BT) by the following scheme.

**Algorithm 1.** FISTA-BT algorithm.

Initialization: \( t_0 = 1 \) and \( \alpha = 1/L \).

Iterative step: let \( k^0 = k^1 \in H \) and calculate \( k^{n+1} \) as follows:

Step 1. Compute the inertial step:

\[
x^n = k^n + t_n (k^n - k^{n-1}),
\]

(8)

where \( t_n = (1 + \sqrt{1 + 4t_{n-1}^2})/2 \) and \( \theta_n = (t_{n-1} - 1)/t_n. \)

Step 2. Compute the \( k^{n+1} \) step:

\[
k^{n+1} = \text{prox}_{\alpha_n}(x^n - \alpha_n \nabla f(x^n)).
\]

(9)

Set \( n = n + 1 \) and return to Step 1.

Without the Lipschitz condition on the gradient of functions, Cruz and Nghia [22] proposed a new version of the forward-backward method (FISTA-CN) based on the linesearch rule.

**Algorithm 2.** FISTA-CN algorithm.

Initialization: \( t_0 = 1, \sigma > 0, \theta \in (0, 1), \) and \( \delta \in (0, 1/2) \).

Iterative step: let \( k^0 = k^1 \in H \) and calculate \( k^{n+1} \) as follows:

Step 1. Compute the inertial step:

\[
x^n = k^n + t_n (k^n - k^{n-1}),\]

(10)

\[
y^n = P_D(x^n),
\]

where \( t_n = (1 + \sqrt{1 + 4t_{n-1}^2})/2 \) and \( \theta_n = (t_{n-1} - 1)/t_n. \)

Step 2. Compute the \( k^{n+1} \) step:

\[
k^{n+1} = \text{prox}_{\alpha_n}(y^n - \alpha_n \nabla f(y^n)),
\]

(11)

where \( \alpha_n = \sigma \theta^n m_n \) and \( m_n \) is the smallest number such that

\[
\sigma \theta^n \|\nabla f(\text{prox}_{\alpha_n}(y^n - \alpha_n \nabla f(y^n))) - \nabla f(y^n)\| \leq \delta \|\text{prox}_{\alpha_n}(y^n - \alpha_n \nabla f(y^n)) - y^n\|.
\]

(12)

Stop criteria if \( k^{n+1} = y^n \), then stop.

If \( k^{n+1} \neq y^n \), then set \( n = n + 1 \) and return to Step 1.

In 2017, Verma and Shukla [23] introduced the new accelerated proximal gradient algorithm (NAGA) which is generated by the following.

**Algorithm 3.** NAGA algorithm.

Iterative step: let \( k^0 = k^1 \in H \) and calculate \( k^{n+1} \) as follows:

Step 1. Compute the inertial step:

\[
x^n = k^n + t_n (k^n - k^{n-1}).
\]

(13)
Step 2. Compute

\[ y_n = (1 - \alpha_n)x^n + \alpha_n \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)), \]

\[ k^{n+1} = \text{prox}_{\alpha_n g}(y_n - \alpha_n \nabla f(y_n)), \]  

where \( \alpha_n \in (0, 2/L) \). Set \( n = n + 1 \) and return to Step 1.

This work presents a new splitting method called a new modified forward-backward splitting algorithm (NMFBS) for convex minimization problems. Our results extend and improve the corresponding results of Tseng [19] and Cruz and Nghia [22]. The step size defined in this work does not require the Lipschitz condition of the gradient functions. Finally, we also present the numerical experiments of our algorithm for solving image recovery problems and show the comparison of our proposed method to FISTA-BT [21], FISTA-CN [22], and NAGA [23].

2. Main Theorem

We assume that \( f : H \rightarrow \mathbb{R} \cup \{+\infty\} \) and \( g : H \rightarrow \mathbb{R} \cup \{+\infty\} \) are proper, lower semicontinuous, and convex functions; \( f \) is uniformly continuous on bounded sets; and \( \nabla f \) is bounded on bounded sets. The following is our algorithm.

Algorithm 4. The new modified forward-backward splitting algorithm (NMFBS)

Initialization: given \( \sigma > 0, \theta > 0, \delta \in (0, 1), \) and \( \theta_1 > 0 \).

Iterative step: let \( k^0 = k^1 \in H \) and calculate \( k^{n+1} \) as follows:

\[ y_n = (1 - \alpha_n)x^n + \alpha_n \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)), \]

\[ k^{n+1} = \text{prox}_{\alpha_n g}(y_n - \alpha_n \nabla f(y_n)), \]
Step 1. Compute the inertial step:

\[ x^n = k^n + \theta_n (k^n - k^{n-1}). \]  \hfill (15)

Step 2. Compute:

\[ y^n = \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)) + \alpha_n \left( \nabla f(x^n) - \nabla f\left( \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)) \right) \right) \]  \hfill (16)

where \( \alpha_n = \sigma \theta^m_n \) and \( m_n \) is the smallest number such that

\[
\alpha^2_n \left( \left\| \nabla f(x^n) - \nabla f\left( \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)) \right) \right\|^2 + \left\| \nabla f(y^n) - \nabla f\left( \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) \right) \right\|^2 \right) \\
\leq \delta^2 \left( \left\| x^n - \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)) \right\|^2 + \left\| y^n - \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) \right\|^2 \right)
\]  \hfill (17)

Step 3. Compute the \( k^{n+1} \) step:

\[ k^{n+1} = \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(y^n)) + \alpha_n \left( \nabla f(y^n) - \nabla f\left( \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) \right) \right) \]  \hfill (18)

Set \( n = n + 1 \) and return to Step 1.

Following the proof as in [24], we can show the following lemma.

**Lemma 1.** The linesearch (17) has a finite step.

**Theorem 2.** Suppose that \( \alpha_n \geq \alpha \) for some \( \alpha > 0 \), \( \theta_n \geq 0 \), and \( \sum_{n=0}^{\infty} \theta_n < +\infty \). Then, \( (k^n) \) generated by Algorithm 4 converges weakly to a minimizer of \( f + g \).

**Proof.** Let \( k_n \in \text{argmin}(f + g) \), and set \( p^n = \text{prox}_{\alpha_n g}(x^n - \alpha_n \nabla f(x^n)) \). Then, we obtain

\[ y^n = p^n + \alpha_n (\nabla f(x^n) - \nabla f(p^n)). \]  \hfill (19)

Moreover, we have

\[ x^n - p^n - \alpha_n \nabla f(x^n) \in \alpha_n \partial g(p^n). \]  \hfill (20)

Using (19), we see that

\[ \alpha_n \nabla f(x^n) = y^n - p^n + \alpha_n \nabla f(p^n). \]  \hfill (21)
Combining (20) and (21), we have
\[ x^n - y^n - \alpha_n \nabla f(p^n) \in \alpha_n \partial g(p^n). \tag{22} \]

Now, set \( r^n = \text{prox}_{\alpha g}(y^n - \alpha_n \nabla f(y^n)) \). Then, we obtain
\[ k^{n+1} = r^n + \alpha_n (\nabla f(y^n) - \nabla f(r^n)). \tag{23} \]

Also, we have
\[ y^n - k^{n+1} - \alpha_n \nabla f(r^n) \in \alpha_n \partial g(r^n). \tag{24} \]

Since \( k_s \in \text{argmin}(f + g) \), we obtain \(-\alpha_n \nabla f(k_s) \in \alpha_n \partial g(k_s)\). Thus, by (22), (24), and the monotonicity of \( \partial g \), we have
\[
\begin{align*}
\langle x^n - y^n - \alpha_n (\nabla f(p^n) - \nabla f(k_s)), p^n - k_s \rangle & \geq 0, \\
\langle y^n - k^{n+1} - \alpha_n (\nabla f(r^n) - \nabla f(k_s)), r^n - k_s \rangle & \geq 0.
\end{align*} \tag{25}
\]

So, we have \( \langle x^n - y^n, p^n - k_s \rangle \geq 0 \) and \( \langle y^n - k^{n+1}, r^n - k_s \rangle \geq 0 \) by the monotonicity of \( \nabla f \). Thus, we have
\[
\begin{align*}
\langle x^n - y^n, p^n - y^n \rangle + \langle x^n - y^n, y^n - k_s \rangle & \geq 0, \\
\langle y^n - k^{n+1}, r^n - k^{n+1} \rangle + \langle y^n - k^{n+1}, k^{n+1} - k_s \rangle & \geq 0. 
\end{align*} \tag{26, 27}
\]

We note that \( \|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y \rangle + \|y\|^2 \) for all \( x, y \in H \). Using (26), we have
\[
2(\langle x^n - y^n, p^n - y^n \rangle + \langle x^n - y^n, y^n - k_s \rangle) = \|x^n - k_s\|^2 - \|y^n - k_s\|^2 + \|y^n - p^n\|^2 - \|x^n - p^n\|^2. \tag{28}
\]

Using (27), we have
\[
2(\langle y^n - k^{n+1}, r^n - k^{n+1} \rangle + \langle y^n - k^{n+1}, k^{n+1} - k_s \rangle) = \|y^n - k_s\|^2 - \|k^{n+1} - k_s\|^2 + \|k^{n+1} - r^n\|^2 - \|y^n - r^n\|^2. \tag{29}
\]

From (26), (27), (28), and (29), we obtain
\[
\|k^{n+1} - k_s\|^2 \leq \|x^n - k_s\|^2 + \|y^n - p^n\|^2 - \|x^n - p^n\|^2 + \|k^{n+1} - r^n\|^2 - \|y^n - r^n\|^2. \tag{30}
\]
Using (17), (19), (23), and (30), we obtain

\[
\|k^{n+1} - k_*\|^2 \leq \|x^n - k_*\|^2 + \|p^n + \alpha_n(\nabla f(x^n) - \nabla f(p^n)) - p^n\|^2 + \|\nabla f(y^n) - \nabla f(r^n)) - r^n\|^2
\]

\[
\left(\|x^n - \theta_k p^n\|^2 + \|y^n - r^n\|^2\right) - \|x^n - \theta_k p^n\|^2 - \|y^n - r^n\|^2
\]

\[
= ||x^n - k_*||^2 - (1 - \delta^2)||x^n - p^n||^2 - (1 - \delta^2)||y^n - r^n||^2.
\]

(31)

Next, we will show that \(\lim_{n \to \infty} \|k^n - k_*\|\) exists. From (31), we see that

\[
\|k^{n+1} - k_*\| \leq \|x^n - k_*\| + \|p^n + \alpha_n(\nabla f(x^n) - \nabla f(p^n)) - p^n\|^2 + \|\nabla f(y^n) - \nabla f(r^n)) - r^n\|^2
\]

\[
= \|k^n - k_*\| + \|p^n\| + \|\nabla f(y^n) - \nabla f(r^n)) - r^n\|^2
\]

(32)

By Lemma 5 in [1], we have

\[
\|k^{n+1} - k_*\| \leq K \cdot \prod_{j=1}^{n} (1 + 2\theta_j),
\]

(33)

where \(K = \max \{\|k^1 - k_*\|, \|k^2 - k_*\|\}\). Since \(\sum_{n=1}^{\infty} \theta_n < +\infty\), we have \(\sum_{n=1}^{\infty} \theta_n \|k^n - k_*\|^2 < +\infty\). By Lemma 1 in [25] and (32), we have \(\lim_{n \to \infty} \|k^n - k_*\|\) that exists. From (31), we see that

\[
\|k^{n+1} - k_*\|^2 \leq \|x^n - k_*\|^2 - (1 - \delta^2)||x^n - p^n||^2
\]

\[
- (1 - \delta^2)||y^n - r^n||^2
\]

\[
= \|k^n - k_*\|^2 + \|\nabla f(x^n) - \nabla f(p^n)) - p^n\|^2 + \|\nabla f(y^n) - \nabla f(r^n)) - r^n\|^2
\]

\[
+ 2\theta_n \|k^n - k_*\| \|k^n - k_*\| + \theta_n \|k^n - k_*\| \|k^n - k_*\|^2
\]

\[
- (1 - \delta^2)||x^n - p^n||^2 - (1 - \delta^2)||y^n - r^n||^2.
\]

(34)

Noting \(\lim_{n \to \infty} \alpha_n||k^n - k^{n-1}|| = 0\), \(\lim_{n \to \infty} \|k^n - k_*\|\) exists and \(\delta \in (0, 1)\), we have

\[
\lim_{n \to \infty} \|x^n - p^n\| = 0,
\]

(35)

\[
\lim_{n \to \infty} \|y^n - r^n\| = 0.
\]

(36)

Since \(\nabla f\) is uniformly continuous on bounded sets, we have

\[
\lim_{n \to \infty} \|\nabla f(x^n) - \nabla f(p^n))\| = 0,
\]

(37)

\[
\lim_{n \to \infty} \|\nabla f(y^n) - \nabla f(r^n))\| = 0.
\]

(38)

By definition of \(x^n\), it is easy to see that \(\lim_{n \to \infty} \|x^n - k^n\| = 0\). Then,

\[
\|p^n - k^n\| \leq \|x^n - p^n\| + \|x^n - k^n\| \to 0 \text{ as } n \to \infty.
\]

(39)
From (35), (36), (37), and (39), we obtain
\[ ||r^n - k^n|| \leq ||r^n - y^n|| + ||y^n - p^n|| + ||p^n - k^n|| = ||r^n - y^n|| + \alpha_n (\nabla f(x^n) - \nabla f(p^n)) - p^n|| + ||p^n - k^n|| = ||r^n - y^n|| + \alpha_n \nabla f(x^n) - \nabla f(p^n)\| + \|p^n - k^n\| \rightarrow 0 \text{ as } n \rightarrow \infty. \]

(40)

\[
y^n - r^n = \nabla f(y^n) = \frac{y^n - \text{prox}_{\alpha_n g}(y^n - \alpha_n \nabla f(y^n)) - \alpha_n \nabla f(y^n)}{\alpha_n} \in \partial g(y^n - \alpha_n \nabla f(y^n)).
\]

(41)

It follows that
\[
y^n - r^n = \frac{\nabla f(r^n) - \nabla f(y^n)}{\alpha_n} \in \nabla f(r^n) + \partial g(r^n) \subseteq \partial(f + g)(r^n).
\]

(42)

By passing \( i \rightarrow \infty \) and using (36) and (38), we have \( 0 \in \nabla f(\bar{k}) + \partial g(\bar{k}) \) by Fact 2.2 in \([22]\). Hence, by Theorem 5.5 in \([26]\), we can conclude that \( (k^n) \) converges weakly to a point in \( \text{argmin}(f + g) \). We thus complete the proof. \( \square \)

Remark 3. The condition that \( \alpha_n \geq \alpha > 0 \) for some \( \alpha \) can be dropped in case \( \nabla f \) is Lipschitz continuous on \( H \) since it is bounded below from 0 (see Proposition 4.4.4.11 \([22]\)).

Remark 4. In the main theorem, we use the linesearch technique to calculate our step size at each iteration unlike the result of \([4, 5, 16, 17]\). It is worth mentioning here also that choice of the step size in our algorithm does not depend on the Lipschitz condition of the gradient function. Our proposed algorithms can be applied in image recovery which are more applicable than those of \([4, 5, 16, 17]\).

### 3. Numerical Experiments

Medical imaging plays a crucial role in modern medicine and image data which are found in various clinical specialties, for routine diagnostics in X-ray imaging, monitoring intraoperative progress during surgical procedures and guidance and diagnosis in ailing. In practice, the degradations are unavoidable because the medical imaging systems limit the intensity of the incident radiation to protect the patient’s health. So how to improve image quality is a good choice for medical analysis. Image processing mainly consists of image deblurring, image denoising, and image inpainting which is a branch that usually can be employed optimization techniques to solve it.

The image restoration problem can be explained as follows:

\[
b = Ak + w,
\]

(43)

where \( b \in \mathbb{R}^{mx1} \) is the observed image, \( A \in \mathbb{R}^{mxn} \) is the blur-ring matrix, \( k \in \mathbb{R}^{nx1} \) is an original image, and \( w \) is additive noise. To solve problem (43), we aim to approximate the original image by transforming (43) to the following LASSO problem \([27]\):

\[
\min_k \left( \frac{1}{2} \|b - Ak\|_2^2 + \lambda \|k\|_1 \right),
\]

(44)

where \( \|\cdot\|_1 \) is \( \ell_1 \)-norm. In general, (44) can be formulated in a general form by estimating the minimizer of sum of two functions when \( f(k) = 1/2 \|b - Ak\|_2^2 \) and \( g(k) = \lambda \|k\|_1 \). We next present our algorithm (NMFBs) for LASSO problem with \( \lambda = 10^{-7} \) and also compare its efficiency with FISTA-BT \([21]\), FISTA-CN \([22]\), and NAGA \([23]\). All computational experiments were written in Matlab 2020b and performed on a 64-bit MacBook Pro Chip Apple M1 and 8 GB of RAM.

Let \( k \) be the original images size \( 448 \times 2993 \) and \( 386 \times 608 \times 3 \), respectively. These are shown in Figure 1. To measure the quality of restored images, we use the peak signal-to-noise ratio (PSNR) in decibel (dB) \([28]\) and the structural similarity index metric (SSIM) \([29]\). The iteration numbers for all algorithms are \( 1200^{th} \).

All parameters are chosen as in Table 1. The initial points \( k^0 = k^1 \) are vectors of ones with the size of original images for all algorithms. The blurred images are shown in Figures 2–4. The parameter \( \theta_n \) of FISTA-BT, FISTA-CN, and NAGA is defined as in Algorithm 1.

The numerical results are reported in Table 2 and Figures 5–8.

From Table 2, we see that numerical experiments of NMFBs are better than those of FISTA-BT, FISTA-CN, and NAGA in terms of PSNR and SSIM for all blur types.

We next provide some experiments of the recovered images for two cases to illustrate the convergence behavior of all algorithms in comparison. We plot the number of iterations versus PSNR and SSIM in Figures 6 and 8.

### 4. Conclusion

We have introduced the modified forward-backward algorithm for solving the convex minimization problem of the sum of two functions in a real Hilbert space. The proposed
algorithm does not need to compute the Lipschitz constant of the gradient of functions. We have proved that the sequence generated by the algorithm weakly converges to a minimizer under some mild conditions. Our result can be applied effectively to solve image recovery as shown in numerical experiments. The comparative experiments showed that the proposed algorithm has a better efficiency than FISTA-BT [21], FISTA-CN [22], and NAGA [23] in terms of PSNR and SSIM for all blur types.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

K. Kankam was supported by School of Science, University of Phayao, under grant no. PBTSCE65023. W. Cholamjiak was supported by the Thailand Science Research and Innovation Fund and the University of Phayao under grant no. FF65-UoE002 and no. FF65-RIM072. P. Cholamjiak was supported by the National Research Council of Thailand (NRCT) under grant no. N41A640094.

References


