Some recent advances on iteration for nonsmooth convex minimization problems

Koketso Malebana

School of Science and Technology,

Department of Mathematics and Applied Mathematics

SEFAKO MAKGATHO
HEALTH SCIENCES UNIVERSITY

September 7, 2026

Outline

- Introduction
- Preliminaries
- Main Results
- Numerical Example
- Conclusion
- References

Introduction

- Convex minimization problems (CMP) play a central role in various areas of applied mathematics and data science.
- They arise in numerous domains such as image restoration, signal processing, medical data classification, and machine learning.
- Over time, many researchers have proposed and developed a range of iterative methods for solving CMP (see [1, 2, 3, 4, 5, 6]).
- Let $\mathcal{S}$ be a real Hilbert space and $\Psi : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower-semicontinuous and convex function. The CMP is to find $\hat{u} \in \mathcal{S}$ such that

$$0 \in \nabla\Phi(\hat{u}) + \partial\Psi(\hat{u}), \tag{1}$$

- where $\Phi : \mathcal{S} \to \mathbb{R} \cup \{+\infty\}$ is convex and differentiable on $\mathcal{S}$.
- If $\nabla\Phi(\hat{u})$= 0, then problem (1) reduces to finding $\hat{u} \in \mathcal{S}$ such that

$$0 \in \partial\Psi(\hat{u}).$$

- Various methods such as the proximal point method, the forward-backward splitting method, the projected forward-backward splitting method, and the forward-backward method have been developed by many researchers to solve convex minimization problems.
- The problem (1) is equivalent to the following fixed point problem:

$$x^* = (prox_{\gamma\partial\Psi}(I - \gamma\nabla\Phi))x^*. \tag{2}$$

- According to the above fixed point equation, we can immediately get the well-known Proximal Gradient Method (in short, PGM) algorithm for solving (1) as follows:

$$x_{n+1} = (prox_{\gamma_n \partial \Psi}(I - \gamma_n \nabla \Phi))x_n, \tag{3}$$

where $\text{prox}_{\gamma_n \partial \Psi}(x) = argmin_{u \in \mathcal{S}} \left\{ \Psi(u) + \frac{1}{2\gamma_n} \|u - x\|^2 \right\}$ and the stepsize $\{\gamma_n\}$ is a sequence of positive real numbers.

- In 2009, Beck and Teboulle [4] introduced the fast iterative shrinkage-thresholding algorithm for linear inverse problem (**FISTA**). Let $q_0 = 1$ and $s_0 = s_1 \in \mathbb{H}$. Compute

$$\begin{cases} q_{k+1} = \dfrac{1+\sqrt{1+4q_{k-1}^2}}{2} \\ \beta_k = \dfrac{q_{k-1}-1}{q_k}, \\ u_k = s_k + \beta_k(s_k - s_{k-1}), \\ s_{k+1} = prox_{\frac{1}{L}\Psi}(u_k - \dfrac{1}{L}\nabla\Phi(u_k)), k \geq 1. \end{cases} \tag{4}$$

- This inclusion of the inertial term $u_k = s_k + \beta_k(s_k - s_{k-1})$ improves the convergence speed of (3).
- The authors established a weak convergence of their iterative method.

- Guo and Cui [17] proposed the following modified proximal gradient method (**MPGM)** for solving CMP (1) as follows

$$x_{k+1} = \alpha_k f(x_k) + (1-\alpha_k)\, prox_{\lambda_k \Psi}\, (I - \lambda_k \nabla \Phi)(x_k) + e(x_k), \qquad (5)$$

where $\{\alpha_k\} \in [0,1],\ 0 < a = \inf_k\ \lambda_k \leq \lambda_k < \frac{2}{L}$, $f : \mathbb{H} \to \mathbb{H}$ is a contraction and $e : \mathbb{H} \to \mathbb{H}$ is a perturbation operator satisfying $\sum_{k=0}^{\infty} \|e(x_k)\| < +\infty$.

- They obtained a strong convergence theorem for the sequence generated by (12) for approximating solution to (1) in a real Hilbert space.
- Also, in 2000, Tseng [29] proposed a modified forward-backward algorithm (**MFB**) via the stepsizes with linesearch technique as follows

- Given $\varsigma > 0, \varrho \in (0,1), \sigma \in (0,1)$ and x_1 *in* $\mathbb{H}$. Compute

$$\begin{cases} y_k = prox_{\lambda_k \Psi}(x_k - \lambda_k \nabla \Phi(x_k)), \\ x_{k+1} = y_k - \lambda_k(\nabla \Phi(y_k) - \nabla \Phi(x_k)), k \geq 1 \end{cases} \tag{6}$$

  where λ_k is the largest $\lambda \in \{\varsigma, \varsigma\varrho, \varsigma\varrho^2, ...\}$ satisfying $\lambda_k ||\nabla \Phi(y_k) - \nabla \Phi(x_k)|| \leq ||y_k - x_k||$.

- Mokaba *et al.* proposed the following modified forward-backward splitting method (**MFBSM**) for solving CMP (1). Given $\beta_1 > 0$ and $\rho \in (0,1)$. Let $x_0, x_1 \in \mathbb{H}$ be two initial points and set $k = 1$.

- Step 1: Compute

$$\begin{cases} w_k = z_k + \theta(x_k - x_{k-1}), \\ t_k = x_k + \tau_k(x_k - x_{k-1}), \\ y_k = prox_{\lambda_k \Psi}(I - \beta_k \nabla \Phi) t_k. \end{cases}$$

where

$$\beta_{k+1} = \begin{cases} min\left\{ \dfrac{\rho||t_k - y_k||}{\|\nabla\Phi(t_k) - \nabla\Phi(y_k)\|} \right\}, \ if \ \nabla\Phi(t_k) \neq \nabla\Phi(y_k), \\ \beta_k + \lambda_k \quad , otherwise. \end{cases}$$

If $y_k = t_k$, then stop. Otherwise proceed to the next step.

Step 2: Compute

$$z_{k+1} = y_k + \beta_k(\nabla\Phi(t_k) - \nabla\Phi(y_k))$$

Step 3: Compute

$$x_{k+1} = (1 - \eta_k)w_k + \eta_k z_k.$$

- Over the past decades, several results that were initially developed in linear spaces for solving problems such as equilibrium problems, variational inequalities and convex minimization problems have been extended to nonlinear setting, particularly to Hadamard manifolds.
- This shift is primarily due to the fact that certain problems cannot be properly formulated within linear spaces and instead require a manifold structure.
- The study of convex minimization problems in nonlinear spaces have proven to be highly valuable for computing medians and means of trees, which play a crucial role in fields such as computational phylogenetics, diffusion tensor imaging, consensus algorithms, and modeling of airway systems in human lungs and blood vessels.

- Therefore, nonlinear spaces, particularly Hadamard manifold provide a more appropriate framework for solving optimization problems than in linear spaces.
- Let $\mathcal{S}$ be a nonempty, closed geodesic convex subset of a Hadamard manifold $\mathbb{M}$, $\mathcal{T}_u\mathbb{M}$ is the tangent space of $\mathbb{M}$ at $u \in \mathbb{M}$ and $\mathcal{T}\mathbb{M}$ is the tangent bundle of $\mathbb{M}$. Then, we are interested in the following CMP, which is to find $u \in \mathcal{S}$ such that

$$0 \in \nabla\Phi(u) + \partial\Psi(u), \tag{7}$$

where $\nabla\Phi(u) \in \mathcal{T}_u\mathbb{M}$ is the Riemannian gradient of Φ, $\partial\Psi(u) \subset \mathcal{T}_u\mathbb{M}$ is the subdifferential of Ψ at u and zero is the zero vector in the tangent space $\mathcal{T}_u\mathbb{M}$.

- In 2019, Ansari and Babu [7] proposed the following proximal point algorithm (for short, PPA) for solving CMP (7) as follows: choose $u_0 \in \mathbb{M}$ and compute u^{k+1} such that

$$0 \in \partial\Psi(u^{k+1}) + \mathcal{R}_{u^{k+1},u^k}\nabla\Phi(u^k) - \frac{1}{\lambda_k}\exp^{-1}_{u^{k+1}} u^k, \forall\, k \in \mathbb{N} \qquad (8)$$

where Φ is differentiable and Lipschitz continuous, $\lambda_k \in (0, +\infty)$ is the sequence of real numbers, $\mathcal{R}_{z^k,u^k}$ is the parallel transport from $\mathcal{T}_{u^k}\mathbb{M}$ to $\mathcal{T}_{z^k}\mathbb{M}$, and $\exp$ is the exponential map.

- A convergence result was proved for the sequence produced by (8) to a solution of the CMP (7).

Review of CMP on Hadamard manifold

Ferreira and Oliveira [19] are one of the first researchers to introduce the convex minimization problem from linear spaces to nonlinear spaces, particularly the Hadamard manifold.

In 2021, Khammahawong *et al.*[7] proposed Tseng's methods for convex minimization problem in Hadamard manifold as follows:

Algorithm

Initialization: *Choose* $\gamma > 0, l \in (0,1)$ *and* $\lambda \in (0,1)$. *Let* $w_0 \in \mathbb{M}$ *be an initial point.*

Iterative step: *Given* $w_k \in \mathcal{C}$, *calculate* w_{k+1} *by*

Step 1: *Compute* q_k *such that*

$$0 \in \partial\Psi(z_k) + R_{z_k,u_k}\nabla\Phi(u_k) - \frac{1}{\lambda_k}\exp_{z_k}^{-1} u_k,$$

where β_k *is chosen to be the largest* $\beta \in \{\gamma, \gamma l, \gamma l^2, ...\}$ *satisfying*

$$\beta||R_{z_k,u_k}\nabla\Phi(u_k) - \nabla\Phi(z_k)|| \leq \lambda d(u_k, z_k). \tag{9}$$

If $z_k = u_k$ *then stop and* u_k *is a solution to CMP* (10). *Otherwise*

Step 2: *Compute*

$$w_{k+1} = \exp_{z_k}(\beta_k(R_{z_k,u_k}\nabla\Phi(u_k) - \nabla\Phi(z_k))).$$

Algorithm

Initialization: *Choose* $\gamma > 0, l \in (0,1)$ *and* $\lambda \in (0,1)$. *Let* $w_0 \in \mathbb{M}$ *be an initial point.*

Iterative Step: *Given* $w_k \in \mathcal{C}$, *calculate* w_{k+1} *by*

Step 1: *Compute* q_k *such that*

$$0 \in \partial\Psi(z_k) + R_{z_k,u_k}\nabla\Phi(u_k) - \frac{1}{\lambda_k}\exp_{z_k}^{-1} u_k,$$

If $z_k = u_k$ *then stop and* u_k *is a solution to CMP* (10). *Otherwise*

Step 2: *Compute*

$$w_{k+1} = \exp_{z_k}(\beta_k(R_{z_k,u_k}\nabla\Phi(u_k) - \nabla\Phi(z_k))).$$

and

$$\tau_{k+1} = \begin{cases} min\left\{\dfrac{\lambda d(u_k,z_k)}{\|R_{z_k,u_k}\nabla\Phi(u_k) - \nabla\Phi(z_k)\|}, \tau_k\right\} & if\ \|R_{z_k,u_k}\nabla\Phi(u_k) - \nabla\Phi(z_k)\| \neq 0 \\ \tau_k \quad , otherwise. \end{cases}$$

- The beauty of iterative method is not only in approximating solutions of optimization problems but also the rate at which they converge.
- In order to speed up the rate of convergence of iterative methods, several authors have employed the inertia technique.
- It serves as a powerful tool to accelerate the convergence of iterative algorithms, making it a cornerstone in the field of optimization.
- Celebrated for its remarkable ability to enhance the speed at which these iterative methods reach their solutions, this technique was first introduced by the influential mathematician Polyak [8] to tackle the challenges of smooth convex optimization problems (see [9, 10, 11])

- In previous studies, there has been discoveries of using the one step inertia technique for the convergence of iterative methods although it has a limitation. Its limitation is that it may not preserve the Fejer monotonicity of the sequence it generates.
- Recently, some results on multi-step inertial methods for solving certain optimization problems have been studied in (see [13, 14, 15]).
- In [16], Polyak discussed that the numerical iteration method with multi-step inertial extrapolation steps could improve the speed of convergence of such methods for solving optimization problems even though neither the convergence analysis nor the rate of convergence result of such multi-step inertial methods are established in [16].

Preliminaries

Some Important Definitions

Definition

A Hadamard manifold is a complete Riemannian manifold that is simply connected and has nonpositive sectional curvature.

Definition

A function $f : \mathbb{M} \to (-\infty, \infty]$ *is called:*

(i) convex, if for any geodesic ς in M, the composition function $f \circ \varsigma : \mathbb{R} \to (-\infty, +\infty]$ is a convex, that is

$$(f \circ \varsigma)(ta + (1-t)b) \leq t(f \circ \varsigma)(a) + (1-t)(f \circ \varsigma)(b),$$

for any $a, b \in \mathbb{R}$ and $0 \leq t \leq 1$.

(ii) proper, if $D(f) \neq \emptyset$ and for all $p \in D(f)$, we have $f(p) > -\infty$.

(iii) lower-semicontinuous on $\mathbb{M}$, if and only if $epi(f)$ is a closed subset of $\mathbb{M} \times \mathbb{R}$, where $epi(f)$ is the epigraph denoted by $\{(x, \beta) \in \mathbb{M} \times \mathbb{R} : f(x) \leq \beta\}$.

Definition

A vector-field $\Phi \in \mathcal{U}(\mathcal{S})$ is said to be:

(i) *monotone, if for any $x, y \in \mathbb{M}$*

$$\langle \Phi(x), \exp_x^{-1} y \rangle \leq \langle \Phi(y), -\exp_y^{-1} x \rangle;$$

(ii) *$\mathcal{L}$-Lipschitz continuous, if there is $\mathcal{L} > 0$ such that*

$$\|\mathcal{R}_{x,y}\Phi(y) - \Phi(x)\| \leq \mathcal{L} d(x, y) \ \forall x, y \in \mathbb{M}.$$

Definition

A vector-field $\Psi \in \mathcal{X}(\mathcal{S})$ *is said to be:*

(i) *monotone, if for any* $p, q \in D(\Psi)$

$$\langle w, \exp_p^{-1} q \rangle \leq \langle u, -\exp_p^{-1} q \rangle, \forall\, w \in \Psi(p) \text{ and } \forall\, u \in \Psi(q).$$

(ii) *maximal monotone if it is monotone and* $\forall p \in \mathcal{S}$ *and* $w \in \mathcal{T}_p\mathbb{M}$, *the condition*

$$\langle w, \exp_p^{-1} q \rangle \leq \langle u, -\exp_p^{-1} q \rangle,$$

$\forall\, q \in D(\Psi)$ *and* $\forall\, u \in \Psi(q)$ *imply that* $w \in \Psi(p)$.

Main Results

Initialization: Given $\delta \in (0,1)$ $\sigma > 0, l \in (0,1), \alpha_k \geq 0$ and $\eta_k \in [0,1]$.

Iterative Step: Let $u^0, u^1 \in \mathbb{M}$ and compute u^{k+1} as follows:

Step 1: Calculate the inertial step:

$$w^k = \exp_{u^k}(-\alpha_k \exp_{u^k}^{-1} u^{k-1}) \ \ and \ \ z^k = \exp_{w^k}(\eta_k \exp_{w^k}^{-1} u^k). \tag{10}$$

Step 2: Find y^k in $\mathbb{M}$ satisfying

$$0 \in \partial\Psi(y^k) + \mathcal{R}_{y^k,z^k} \nabla\Phi(z^k) - \frac{1}{\beta_k} \exp_{y^k}^{-1} z^k, \tag{11}$$

where $\beta_k = \sigma l^{n_k}$ and n_k is the smallest positive number such that:

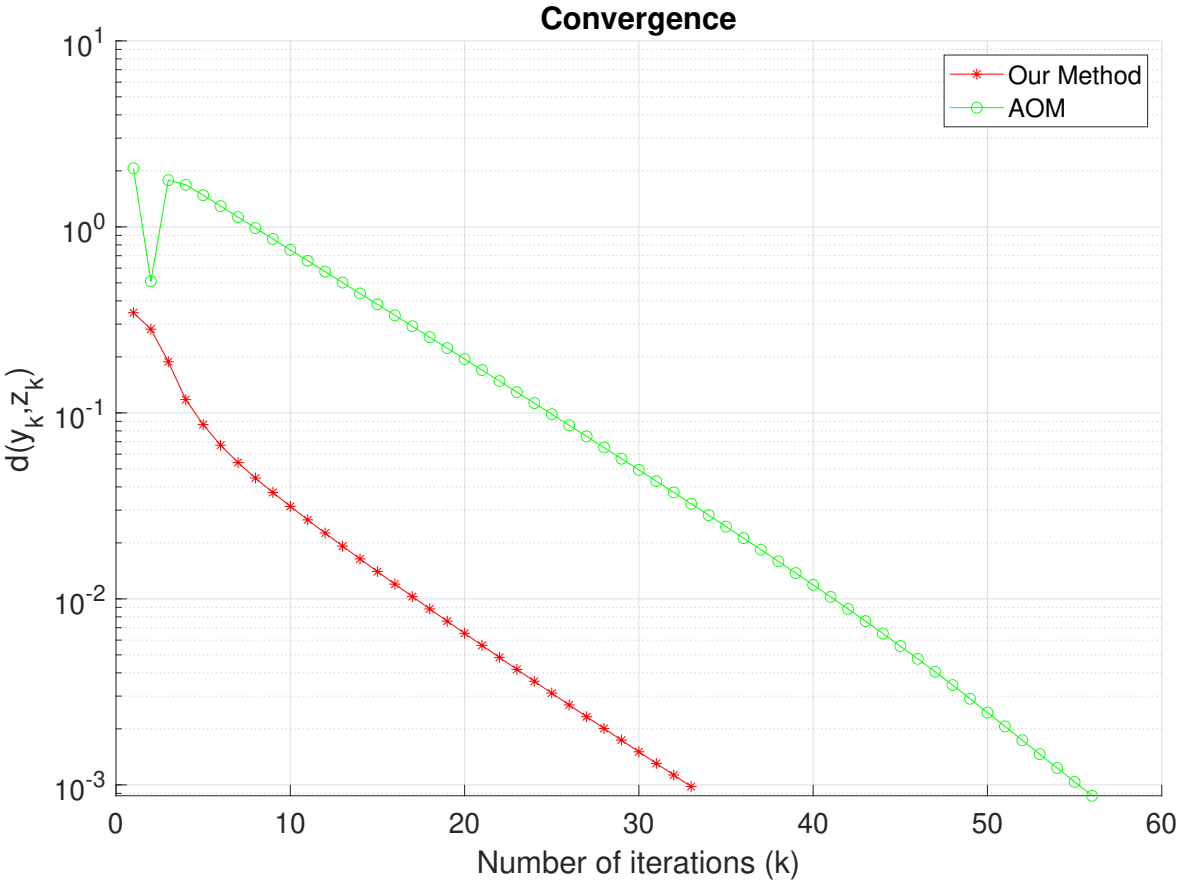
$$\beta_k \|\nabla\Phi(y^k) - R_{y^k,z^k} \nabla\Phi(z^k)\| \leq \delta d(y^k, z^k). \tag{12}$$

Step 3: Compute the u^{k+1} step

$$u^{k+1} = \exp_{y^k}(\beta_k(R_{y^k,z^k} \nabla\Phi(z^k) - \nabla\Phi(y^k))). \tag{13}$$

Set $k := k + 1$ and return to Step 1.

Numerical Example



Conclusion

In this paper, we introduced an alternating modified proximal point algorithm together with a dynamic step size, which incorporates the self-adaptive step size and Armijo linesearch to approximate the solution of monotone inclusion problem in Hadamard manifold. We established a convergence result of the proposed method under the condition that one of the vector fields is monotone and $L-$ Lipschitz continuous.



Cholamjiak, P., Van Hieu, D., and Cho, Y. T. (2021). Relaxed forward-backward splitting methods for solving variational inclusions and applications. Journal of Scientific Computing, 88(3), 85.



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, 1-17.



Cholamjiak, W., Yambangwai D., and Das, S. (2025). Inertial Three-Step Forward-Backward Splitting Algorithms to solve Inclusion Problems and Their Application to Image Restoration Problems. In Mathematical Analysis for Engineering and Applied Sciences. (pp. 141-172). CRC Press.


Hanjing, A., and Suantai, S. (2023). Novel algorithms with inertial techniques for solving constrained convex minimization problems and applications to image inpainting. Mathematics, 11(8), 1813.


Kankam, K., Noor, M. A., and Cholamjiak, P. (2025). Three-step projected forward-backward algorithms for constrained minimization problem. Journal of Applied Mathematics and Computing, 71(1), 465-487.


Suantai, S., Noor, M. A., Kankam, K., and Cholamjiak, P. (2021). Novel forward-backward algorithms for optimization and applications to compressive sensing and image inpainting. Advances in Difference Equations, 2021(1), 265.


Ansari, Q. H., and Babu, F. Proximal point algorithm for inclusion problems in Hadamard manifolds with applications. Optim Lett. 2021;15(3):901-921 .

- Polyak, B. T. Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. Phys.* **1964,** 4, 1-17.
- Attouch, H., and Cabot, A. Convergence rate of a relaxed inertial proximal algorithm for convex minimization. *Optimization* **2019**, 69, 1281-1312.
- Alvarez, F., and Attouch, H. An inertial proximal method for maxi mal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Anal.* bf 2001, 9, 3-11.
- Van Hieu, D. An inertial-like proximal algorithm for equilibrium problems. *Math. Methods Oper. Res.* **2008,** 88, 399-415.
- *B. Martinet, Regularisation d' inequations varaiationnelles par approximations successives, Rev. Fr. Inform.Rec. Oper., **4** (1970), 223-236.*


Dong, L.Q., Cho, J.Y., Zhong, L.L., Rassias, M.Th.: Inertial projection and contraction algorithms for variational inequalities. J. Glob. Optim. **70**, 687-704 (2018).


Thong, D.V., Hieu, D.V., Rassias, T.M.: Self adaptive inertial subgradient extragradient algorithms for solving pseudomonotone variational inequality problems. Optimization Lett. **14**, 115-144 (2020).


Yang, J., Liu, H., Liu, Z.: Modified subgradient extragradient algorithms for solving monotone variational inequalities. Optimization. **67**, 2247-2258 (2018).


Polyak, B.T.: Introduction to optimization. New York: Optimization Software, Publications Division,1987.


Guo, Y., Cui, W., Strong convergence and bounded perturbation resilence of a modified proximal gradient algorithm, J. Ineq. Appl. **2018** (2018).


Mainge, P.E., Approximation methods for common fixed points of nonexpansive mappings in Hilbert spaces, J. Math. Anal. Appl. **325** (2007), 469-479.


Ferreira, O. P., Oliveira, P. R. (2002). Proximal Point Algorithm On Riemannian Manifolds. Optimization, 51(2), 257â270. https://doi.org/10.1080/02331930290019413


Meir, A., Keeler, E., A theorem on contraction mappings, J. Math. Anal. Appl. **28** (1969), 326-329.


Tseng, P.: A modified foward-backward splitting method for maximal monotone mappings. SIAM J. Control. Optim. **38**, 431-446 (2000).


Chen. X and Duan. D, Inertial self-adaptive algorithms for solving non-smooth convex optimization problems, Number. Algor. (2024), 1-31.


Auslender, A., Teboulle, M., Ben-Tiba, S.: A logarithmic-quadratic proximal method for variational inequalities. Comput. Optim. Appl. Optim. **12**, 31-40 (1999).


Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics. Springer, New York (2011).


S. Suantai and P. Thongsri, A Fast Forward-Backward Algorithm Using Linesearch and Inertial Techniques for Convex Bi-level Optimization Problems with Applications, *Capathian J. Math.,* **40(3)**, 769-788 (2024)