

RELAXED INERTIAL PROJECTIVE
FORWARD-BACKWARD SPLITTING ALGORITHMS
FOR REGULARIZED LEAST SQUARE PROBLEMS

Watcharaporn Cholanjiak¹, Hemen Dutta^{2✉}

Received on December 26, 2023

Presented by S. Margenov, Corresponding Member of BAS, on February 27, 2024

Abstract

This study presents flexible conditions of the inertial extrapolation parameter for easy implementation that is added to the algorithm in faster convergence. Modified inertial forward-backward splitting algorithms for solving variational inclusion problems are introduced to apply to solve the LASSO problem for image restoration and signal recovery, and the elastic net model for classification problem. Projection methods are used in the final step for narrowing down the search field, resulting in a better solution.

Key words: inertial technique, projective method, regularized least square problem, forward-backward splitting algorithm, variational inclusion problems

2020 Mathematics Subject Classification: 46T99, 47B02, 47H10, 47J25, 49M05

1. Introduction. Regularized least square problems (RLSB) have a variety of specific applications. In this paper, we consider three applications of the RLSB that are defined as follows:

- (i) image restoration

$$(1.1) \quad \min_x \left(\frac{1}{2} \|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda \|x\|_1 \right)$$

The authors would like to thank the National Research Council of Thailand and University of Phayao (N42A650334), and the University of Phayao and Thailand Science Research and Innovation Fund (Fundamental Fund 2024).

<https://doi.org/10.7546/CRABS.2024.04.03>

which is the process of recovering a degraded image $\mathcal{B} = \mathcal{H}x - \epsilon$, where $x \in \mathbb{R}^n$ is an original image, ϵ is additive noise, $\mathcal{H} \in \mathbb{R}^{m \times n}$ is the blurring matrix, and $\lambda \geq 0$ is a regularized parameter with $\|\cdot\|_1$;

(ii) signal recovery

$$(1.2) \quad \min_x \left(\frac{1}{2} \|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda \|x\|_1 \right)$$

which is the model of recovering a disrupted signal $\mathcal{B} = \mathcal{H}x - \epsilon$, where $x \in \mathbb{R}^N$ is a vector with m nonzero components to be recovered, $\mathcal{B} \in \mathbb{R}^M$ is the observed or measured data with noise ϵ , $\mathcal{H}: \mathbb{R}^N \rightarrow \mathbb{R}^M$ ($M < N$) is a (bounded) linear operator, and $\lambda \geq 0$ is a regularized parameter;

(iii) data classification by extreme learning machine (ELM [5])

$$(1.3) \quad \min_x \left(\frac{1}{2} \|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda_1 \|x\|_1 + \lambda_2 \|x\|_2^2 \right),$$

where x is the optimal output weight, $\mathcal{H}$ is the hidden layer output matrix, $\mathcal{B}$ is the training target data, and $\lambda_1, \lambda_2 \geq 0$ are regularized parameters.

The regularized least square problems (1.1)–(1.2) are called least absolute shrinkage and selection operator (LASSO). Problem (1.3) is called elastic net problem.

Many problems in optimization can be formulated to solved problems (1.1)–(1.3), see [1, 6, 7, 10, 12, 18]. We focus on variational inclusion problems (VIP) which are defined in a real Hilbert space H : to find an element $\hat{x} \in H$ such that

$$(1.4) \quad 0 \in (F + G)\hat{x},$$

where $G: H \rightarrow 2^H$ is a multi-valued maximal monotone mapping and $F: H \rightarrow H$ is a monotone and Lipschitz continuous mapping. The set of solutions of the VIP (1.4) is denoted by $(F + G)^{-1}(0)$. By setting $F(x) \equiv \nabla(\frac{1}{2}\|\mathcal{B} - \mathcal{H}x\|_2^2)$ for problems (1.1)–(1.2) or $F(x) \equiv \nabla(\frac{1}{2}\|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda_2\|x\|_2^2)$ for problem (1.3) and $G(x) \equiv \partial(\lambda\|x\|_1)$ or $G(x) \equiv \partial(\lambda_1\|x\|_1)$, it is well known that VIP (1.4) can be reduced to solve problems (1.1)–1.3.

To solve VIP (1.4) many algorithms [2, 4, 13] were proposed via the resolvent mapping $J_\tau^G: H \rightarrow H$ which is defined by $J_\tau^G(x) := (I + \tau G)^{-1}(x)$ for all $x \in H$ since VIP (1.4) is equivalent to the fixed point equation $\hat{x} = J_\tau^G(I - \tau F)\hat{x}$. The famous one is called the forward-backward splitting algorithm [3, 8] which is defined by x_1 and

$$(1.5) \quad x_{n+1} = J_{\tau_n}^G(x_n - \tau_n Fx_n), \quad n \geq 1,$$

where $\tau_n > 0$ is the stepsize. Algorithmic convergence theorem and convergence acceleration techniques have been of great interest to many mathematicians. The well-known inertial technique for speeding-up the convergence was first introduced by POLYAK [14] which initiated from the heavy ball method. The different two inertial forward-backward splitting algorithms are the basic algorithms of adding the inertial technique which combines the inertial technique with the forward-backward splitting algorithm proposed by MOUDAFI and OLINY [11] in 2003, and LORENZ and POCK [9] in 2015. The inertial forward-backward splitting algorithm of Moudafi and Oliny [11] is in the form:

$$(1.6) \quad \begin{aligned} y_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} &= J_{\tau_n}^G(y_n - \tau_n F x_n), \quad n \geq 1, \end{aligned}$$

where $F: H \rightarrow H$ and $G: H \rightarrow 2^H$. They obtained the weak convergence theorem provided $\tau_n < 2/L$ with L the Lipschitz constant of F and the condition on the inertial extrapolation α_n . The inertial forward-backward splitting algorithm of Lorenz and Pock [9] is in the form:

$$(1.7) \quad \begin{aligned} y_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} &= J_{\tau_n}^G(y_n - \tau_n F y_n), \quad n \geq 1, \end{aligned}$$

where $\{\tau_n\}$ is a positive real sequence. It is observed that algorithms (1.6) and (1.7) are different since operator F in algorithm (1.6) is still evaluated at the point x_n .

In this paper, we discuss the conditions of inertial extrapolation α_n of modified inertial forward-backward splitting algorithms for easy implementation used in three applications (1.1)–(1.3). We also introduce two methods for modifying new inertial forward-backward splitting algorithms and show their advantages. The first one is adding the projection method in the last step to get the next iteration and the second is combining with nonexpansive mapping to approximate a common solution of VIP (1.4) and fixed point problem.

2. Algorithms. In this research, we focus on inertial technique and projective method for modifying new algorithms to solve VIP (1.4). Figure 1 shows the advantage of relax condition of the inertial extrapolation α_1 and adding the metric projection P_C to get x_2 , where C is nonempty closed convex subset of $\mathbb{R}^2$.

Remark 2.1. From Figure 1, we see that

- (i) x_1 (black vector) is updated to new x_1 (blue vector if $\alpha_1 > 0$ and red vector if $\alpha_1 < 0$). This shows that if $\alpha_n \in (-\infty, +\infty)$, then finding acceleration sequence $\{x_n\}$ is more general.
- (ii) x_2 (green vector) is updated by P_C to new x_2 (orange vector if $C = \{x: \|x\|_1 \leq \beta\}$, blue vector if $C = \{x: \|x\|_2^2 \leq \beta\}$, and light blue vector if $C = \{x: \|x\|_\infty \leq \beta\}$).

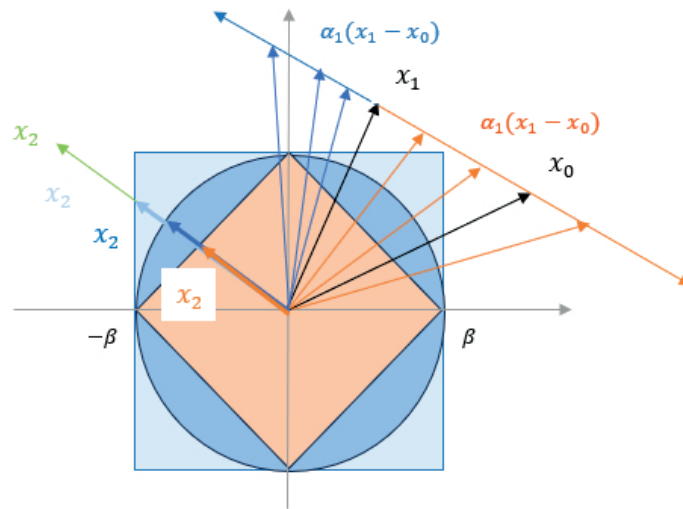


Fig. 1. The structure of an inertial projective method

By Remark 2.1, we present two projection inertial forward-backward splitting algorithms modifying those of Moudafi and Oliny [11], and Lorenz and Pock [9], respectively, as follows:

$$(2.1) \quad \begin{aligned} y_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} &= P_C(J_{\tau_n}^G(y_n - \tau_n F x_n)), \quad n \geq 1, \end{aligned}$$

$$(2.2) \quad \begin{aligned} y_n &= x_n + \alpha_n(x_n - x_{n-1}), \\ x_{n+1} &= P_C(J_{\tau_n}^G(y_n - \tau_n F y_n)), \quad n \geq 1, \end{aligned}$$

with the inertial extrapolation $\alpha_n \in (-\infty, +\infty)$.

Remark 2.2. From algorithms (2.1) and (2.2), we see that

- (i) if $C = H$, algorithms (2.1) and (2.2) reduce to algorithms of Moudafi and Oliny [11], and Lorenz and Pock [9], respectively, but α_n is chosen more general;
- (ii) algorithms (2.1) and (2.2) converge to a solution in $(F + G)^{-1}(0) \cap C$;
- (iii) algorithms (2.1) and (2.2) are applied to solve three applications (1.1)–(1.3) which is more general on constrained set C , where C is nonempty closed and convex subset of H . Then LASSO (1.1)–(1.2) and the elastic net problem (1.3) can be considered in constrained set C , respectively, as follows:

$$(2.3) \quad \min_{x \in C} \left(\frac{1}{2} \|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda \|x\|_1 \right)$$

and

$$(2.4) \quad \min_{x \in C} \left(\frac{1}{2} \|\mathcal{B} - \mathcal{H}x\|_2^2 + \lambda_1 \|x\|_1 + \lambda_2 \|x\|_2^2 \right);$$

- (iv) we can modify a new projection inertial algorithm by modifying the projection P_C in the last step and relax the inertial extrapolation parameter α_n .

Also for algorithms which solve a common solution of VIP (1.4) and fixed point problem of nonexpansive mapping T (that is $T: H \rightarrow H$ such that $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in H$) are useful to apply to solve three applications (1.1)–(1.3) with constrained set C . By setting $T \equiv P_C$, then $C = F(T)$, where $F(T)$ is the set of fixed points of T . In recent years, approximating a common solution of VIP (1.4) and fixed point problem have been proposed, see in [15–17], its applications by considering on constrained set C have not appeared.

We next give an example in Euclidean space E^3 for solving the following minimization:

$$(2.5) \quad \min_{x \in C} \left(\frac{1}{2} (\|x - 4\|^2) + \lambda_1 \|x\|_1 + \lambda_2 \|x\|_2^2 \right).$$

Example 2.3. Let $F(x) = \nabla \left(\frac{1}{2} (\|x - 4\|^2) \right)$, $G(x) = \partial (\lambda_1 \|x\|_1 + \lambda_2 \|x\|_2^2)$ and $C = C_1 = \{x: \|x\|_1 \leq 3\}$. Applying algorithm (2.2) to solve problem (2.5), Table 1 shows numerical results in different cases.

Remark 2.4. From Table 1 and Fig. 2, 3, 4, we see that

- (i) the extrapolation parameter α_n affects the convergence of the algorithm; the inertial algorithm ($\alpha_n \neq 0$) converges faster than the standard algorithm ($\alpha_n = 0$) in the cases of $\lambda_1 = \lambda_2 = 0$ and $\lambda_1 = 3, \lambda_2 = 0$; conversely, in the case $\lambda_1 = 0, \lambda_2 = 5/3$.
- (ii) it is clear that the metric projection P_C also affects the algorithm's convergence; choosing the right C for the algorithm speeds up the convergence.

T a b l e 1

Numerical results of algorithm (2.2) to solve problem (2.5) for different cases

Parameters		$C = H$		$C = C_1$	
		$\alpha_n = 0$	$\alpha_n \neq 0$	$\alpha_n = 0$	$\alpha_n \neq 0$
$\lambda_1 = \lambda_2 = 0$	Number of iterations	62	46	55	41
	CPU times	0.0125	0.0196	0.0388	0.0271
$\lambda_1 = 3, \lambda_2 = 0$	Number of iterations	6	6	6	5
	CPU times	0.0565	0.0308	0.0504	0.0505
$\lambda_1 = 0, \lambda_2 = 5/3$	Number of iterations	23	25	23	25
	CPU times	0.0160	0.0248	0.0265	0.0293

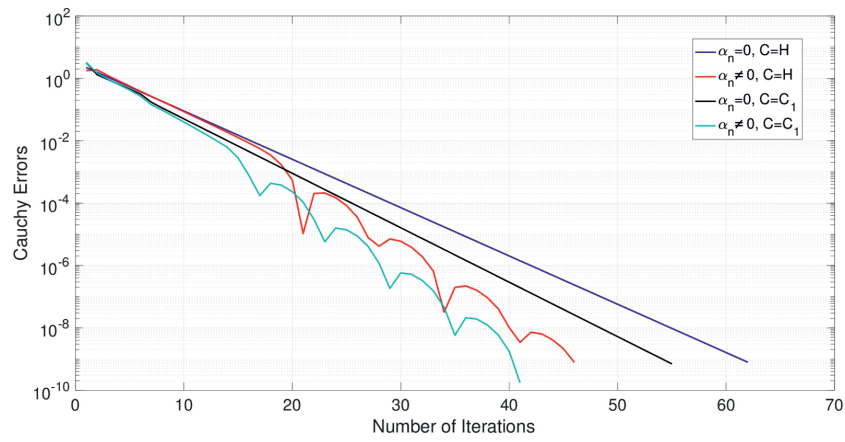


Fig. 2. Cauchy errors plots for $\lambda_1 = \lambda_2 = 0$

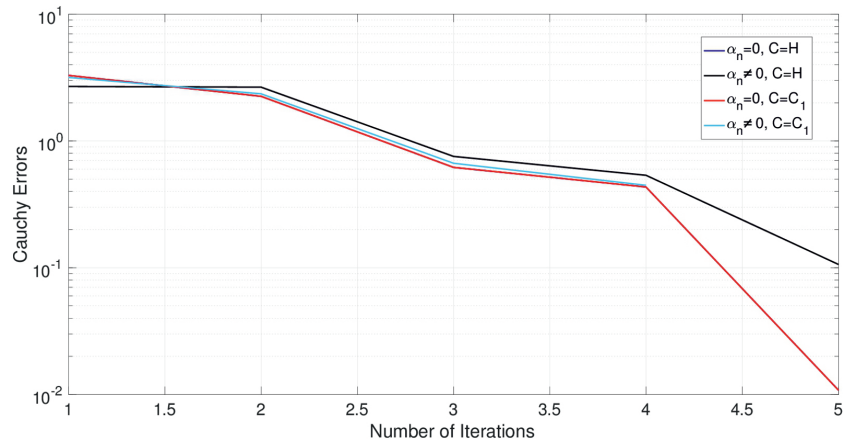


Fig. 3. Cauchy errors plots for $\lambda_1 = 3, \lambda_2 = 0$

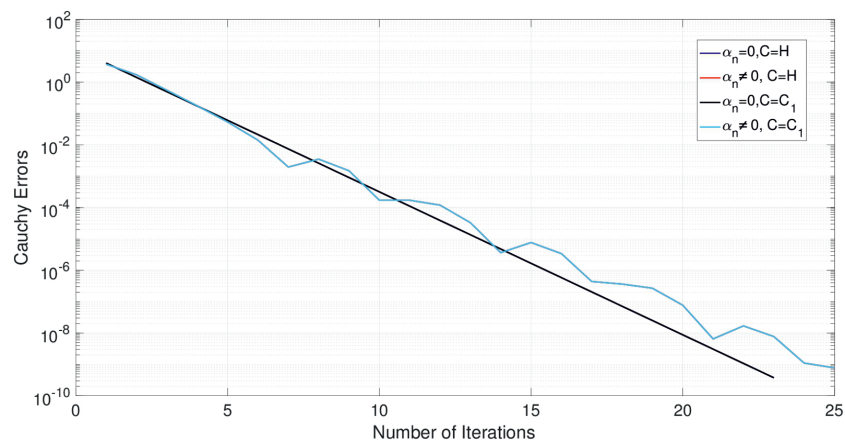


Fig. 4. Cauchy errors plots for $\lambda_1 = 0, \lambda_2 = 5/3$

3. Conclusion. This paper studies the concept of speeding up the convergence techniques of algorithms that solve variational inclusion problems to apply in image restoration, signal recovery, and machine learning. The numerical example shows that adding the inertial term and metric projection are two good ways to improve convergence.

REFERENCES

- [1] AMIR T., R. BASRI, B. NADLER (2021) The trimmed lasso: Sparse recovery guarantees and practical optimization by the generalized soft-min penalty, *SIAM J. Math. Data Sci.*, **3**(3), 900–929.
- [2] CHOLAMJIAK P., D. VAN HIEU, Y. J. CHO (2021) Relaxed forward-backward splitting methods for solving variational inclusions and applications, *J. Sci. Comput.*, **88**(3), 85.
- [3] COMBETTES P. L., V. R. WAJS (2005) Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.*, **4**(4), 1168–1200.
- [4] GIBALI A., D. V. THONG (2018) Tseng type methods for solving inclusion problems and its applications, *Calcolo*, **55**(4), 1–22.
- [5] HUANG G.-B., Q.-Y. ZHU, C.-K. SIEW (2006) Extreme learning machine: Theory and applications, *Neurocomputing*, **70**(1–3), 489–501.
- [6] IBRAHIM A. H., P. KUMAM, M. SUN et al. (2022) Projection method with inertial step for nonlinear equations: application to signal recovery, *J. Ind. Manag. Optim.*, **19**(1), 30–55.
- [7] KANKAM K., N. PHOLASA, P. CHOLAMJIAK (2019) Hybrid forward-backward algorithms using linesearch rule for minimization problem, *Thai J. Math.*, **17**(3), 607–625.
- [8] LÓPEZ G., V. MARTÍN-MÁRQUEZ, F. WANG, H.-K. XU (2012) Forward-backward splitting methods for accretive operators in Banach spaces, *Abstr. Appl. Anal.*, **2012**, Art. ID 109236.
- [9] LORENZ D., T. POCK (2015) An inertial forward-backward algorithm for monotone inclusions, *J. Math. Imaging Vis.*, **51**, 311–325.
- [10] MESSAI M., A. AISSA-EL-BEY, K. AMIS, F. GUILLOUD (2019) Iteratively reweighted two-stage LASSO for block-sparse signal recovery under finite-alphabet constraints, *Signal Processing*, **157**, 73–77.
- [11] MOUDAFI A., M. OLINY (2003) Convergence of a splitting inertial proximal method for monotone operators, *J. Comput. Appl. Math.*, **155**, 447–454.
- [12] PADCHAROEN A., P. KUMAM (2020) Fixed point optimization method for image restoration, *Thai J. Math.*, **18**(3), 1581–1596.
- [13] PEEYADA P., R. SUPARATULATORN, W. CHOLAMJIAK (2022) An inertial Mann forward-backward splitting algorithm of variational inclusion problems and its applications, *Chaos, Solitons & Fractals*, **158**, 112048.
- [14] POLYAK B. T. (1964) Some methods of speeding up the convergence of iteration methods, *U.S.S.R. Comput. Math. Math. Phys.*, **4**(5), 1–17 (in Russian).
- [15] SHEHU Y., F. U. OGBUISI (2016) An iterative method for solving split monotone variational inclusion and fixed point problems, *Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM*, **110**, 503–518.

- [16] SUNTHRAYUTH P., P. CHOLAMJIAK (2021) A modified extragradient method for variational inclusion and fixed point problems in Banach spaces, *Applicable Anal.*, **100**(10), 2049–2068.
- [17] ZHANG S. S., J. H. LEE, C. K. CHAN (2008) Algorithms of common solutions to quasi variational inclusion and fixed point problems, *Appl. Math. Mec.*, **29**(5), 571–581.
- [18] ZHANG Y., Y. DAI, Q. WU (2022) An accelerated optimization algorithm for the elastic-net extreme learning machine, *Int. J. Mach. Learn. Cybern.*, **13**(12), 3993–4011.

¹*School of Science, University of Phayao, Phayao 56000, Thailand*
e-mail: watcharaporn.ch@up.ac.th

²*Department of Mathematics, Gauhati University, Guwahati-781014, India*
e-mail: hemen_dutta08@rediffmail.com