Thai Journal of Mathematics
Volume 21 Number 2 (2023)
Pages 351–365

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

International Conference on Digital Image Processing and Machine Learning 2022

New Iterative Method with Inertial Technique for Split Variational Inclusion Problem to Classify TPACK Level of Pre-Service Mathematics Teachers

Suparat Kesornprom[1], Pronpat Peeyada[1], Watcharaporn Cholamjiak[1], Thanyaluck Ngamkhum[2] and Nipa Jun-on[2,*]

[1] *School of Science, University of Phayao, Phayao 56000, Thailand*
e-mail : suparat.ke@gmail.com (S. Kesornprom); pronpat.pee@gmail.com (P. Peeyada); watcharaporn.ch@up.ac.th (W. Cholamjiak)

[2] *Department of Mathematics, Faculty of Science, Lampang Rajabhat University, Lampang 52100, Thailand*
e-mail : thanyaluck@g.lpru.ac.th (T. Ngamkhum); nipa.676@g.lpru.ac.th (N. Jun-on)

Abstract The study proposes new iterative algorithms employing inertial technique and self-adaptive stepsize for solving a split variational inclusion problem in a Hilbert space and providing a weak convergence theorem for the proposed algorithm. Finally, we present several numerical experiments to solve the problem of data classification to classify pre-service mathematics teachers' TPACK levels through self-assessment-based measures as its application.

MSC: 49J40; 47J20; 35A15; 65K10; 97C70; 97U70
Keywords: variational inclusion problem; self adaptive; inertial method; weak convergence; data classification; TPACK

Submission date: 25.11.2022 / Acceptance date: 20.04.2023

1. Introduction

Shulman [25, 26] introduced the concept of Pedagogical Content Knowledge (PCK) to the attention of the education world. Despite the fact that Shulman's original conception of PCK contained technology, the rapid infusion of technology into our society and education, particularly digital technology, necessitated a more explicit augmentation of the concept with technology. This led to the invention of the phrases information and communication technology (ICT)-related PCK and technologically-enhanced PCK (Angeli Valanides, 2005; Niess, 2005). This was then developed into the Technology, Pedagogy, and Content Knowledge conceptual framework. Initially abbreviated as TPCK (Koehler

*Corresponding author.

Published by The Mathematical Association of Thailand.

Mishra, 2005), this paradigm was renamed TPACK (i.e., entire package) to better emphasize the interdependence of the contributing knowledge fields (Koehler Mishra, 2008; Koehler, Mishra, Kereluik, Shin, Graham, 2014; Thompson Mishra, 2007). TPACK originally included Technological Knowledge (TK), Pedagogical Knowledge (PK), and Content Knowledge (CK). From their intersections, other blended knowledge domains such as TPK, TCK, PCK, and TPACK can be derived.

Numerous studies of TPACK framework have been done to investigate the links between its constructs and to comprehend its measurement. In this study, the new iterative method with inertial technique was proposed to classify the TPACK level of pre-service mathematics teachers from various domains defined as analyzed attributes, while class of the TPACK level was obtained from self-assessment-based measures. We start proposing new iterative method with inertial technique for split variational inclusion problem as the following.

Let H be a real Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. Let $B : H \to 2^H$ be a set-valued maximal monotone mapping. Then the resolvent mapping $J_\beta^B : H \to H$ associated with B is defined by

$$J_\beta^B(x) = (I + \beta B)^{-1}(x), \forall x \in H, \tag{1.1}$$

for some $\beta > 0$, where I stands for the identity operator on H.

We investigate the following split variational inclusion problem (SVIP) which is finding $x^* \in H_1$ such that

$$0 \in B_1(x^*) \text{ and } 0 \in B_2(Ax^*), \tag{1.2}$$

where $B_1 : H_1 \to 2^{H_1}$ and $B_2 : H_2 \to 2^{H_2}$ are multi-valued maximal monotone mappings, $A : H_1 \to H_2$ is a linear and bounded operator.

It is well known that split variational inclusion problem serves as a unified model for many problems of fundamental importance, including the split common fixed-point problem, split variational inequality problem, split zero problem and split feasibility problem [1, 5, 6, 9–11, 18, 23, 30]. Moreover, the split variational inclusion problem has been applied to solving many real life problems, such as in signal processing, approximating theory, data compression (see, e.g., [7, 13, 14, 17, 20, 31]).

Very recently, Byrne et al. [8] studied an iterative method for split variational inclusion problem. Given $x_1 \in H_1$ and $\{x_n\}$ is a sequence defined as follows:

$$x_{n+1} = J_\beta^{B_1}(x_n - \lambda A^*(J_\beta^{B_2} - I)Ax_n), \forall n \geq 1, \tag{1.3}$$

where A^* is the adjoint of A, $\lambda \in (0, 2/L)$ and L is the spectral radius of the operator A^*A. They obtained weak and strong convergence theorems in Hilbert spaces.

In 2001, Alvarez and Attouch [2] applied the inertial technique to obtain an inertial proximal method, which is as follows: let $x_{n-1}, x_1 \in H$, $\beta_n > 0$ and $\theta_n \in [0, 1)$. Define the sequence $\{x_n\}$

$$x_{n+1} = J_\beta^{B_1}(x_n + \theta_n(x_n - x_{n-1})), \forall n \geq 1, \tag{1.4}$$

and the term $\theta_n(x_n - x_{n-1})$ is inertial term which and it can be improved rate of convergence (see[2, 24]).

In this paper, motivated by the work of Alvarez and Attouch [2], we introduce an iterative method with inertial term and using the self adaptive stepsize for solving split variational inclusion problem. Weak convergence theorem is established in the framework of Hilbert spaces. Finally, we apply our algorithm to data classification problem to support the implementation of the proposed.

2. Preliminaries and Lemmas

In this section, we provide some basic definitions and lemmas which will be used in the sequel. Let H be a real Hilbert space. In what follows, we use the following notations:
- the symbols $\rightharpoonup$ stands for the weak convergence.
- the symbols $\to$ stands for the strong convergence.

Recall that a mapping $T : H \to H$ is said to be

(1) nonexpansive if

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

(2) firmly-nonexpansive if

$$\langle Tx - Ty, x - y\rangle \geq \|Tx - Ty\|^2, \ \forall x, y \in H.$$

We note that if T is firmly-nonexpansive, then $I - T$ is also firmly-nonexpansive.

(3) L-Lipschitz continuous, if there exists a constant $L > 0$ such that

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

A set-valued mapping $B : H \to 2^H$ is called monotone if for all $x, y \in H$

$$\langle u - v, x - y\rangle \geq 0, \ \forall u \in Bx \text{ and } v \in By.$$

For a set-valued mapping B, graph(B) is defined as graph$(B) := \{(x, u) \in H \times H : u \in B(x)\}$. A monotone mapping $B : H \to 2^H$ is said to be maximal if the graph(B) is not properly contained in the graph of any other monotone mapping. Let $B : H \to 2^H$ be a set-valued maximal monotone operator. The resolvent operator $J^B_\beta : H \to H$ associated with B is defined by

$$J^B_\beta(x) = (I + \beta B)^{-1}(x), \ \forall x \in H,$$

where $\beta > 0$. It is well known that the resolvent operator is single-valued and firmly non-expansive.

Lemma 2.1. (Demiclosedness principle [15]) *Let C be a nonempty closed convex subset of a real Hilbert space H and let $T : C \to C$ be a nonexpansive mapping. If $x_n \rightharpoonup x \in C$ and $\lim_{n\to\infty} \|x_n - Tx_n\| = 0$, then $x = Tx$.*

In order to study the SVIP, we recall some lemmas which are needed in our proof. We denote by $B^{-1}(0) = \{x \in H : 0 \in Bx\}$, $\mathcal{D}(T)$ the domain of T and $Fix(T)$ the fixed point set of T, that is, $Fix(T) = \{x \in H : x = Tx\}$.

Lemma 2.2. [4] *Let H be a real Hilbert space, $B : H \to 2^H$ be a set-valued maximal monotone mapping for each $x, y \in H$, each $z \in B^{-1}(0)$, and each $\beta > 0$; we have*

(i) $\langle x - J^B_\beta x, z - J^B_\beta x\rangle \leq 0$

(ii) $\|J^B_\beta x - J^B_\beta y\|^2 \leq \langle J^B_\beta x - J^B_\beta y, x - y\rangle$

(iii) $\|J^B_\beta x - z\|^2 \leq \|x - z\|^2 - \|J^B_\beta x - x\|^2$.

Lemma 2.3. [12, 19] *Let H be a real Hilbert space, $B : H \to 2^H$ be a set-valued maximal monotone mapping. Thus,*

(i) J^B_β is a single-valued and firmly nonexpansive mapping for each $\beta > 0$;

(ii) $\mathcal{D}(J^B_\beta) = H$ and $Fix(J^B_\beta) = \{x \in \mathcal{D}(B) : 0 \in Bx\}$;

(iii) $\|x - J_\beta^B x\| \leq \|x - J_\gamma^B x\|$ *for all* $0 < \beta \leq \gamma$ *and for all* $x \in H$*;*

(iv) Suppose that $B^{-1}(0) \neq \emptyset$. *Then* $\|x - J_\beta^B x\|^2 + \|J_\beta^B - x^*\|^2 \leq \|x - x^*\|^2$ *for each* $x \in H$*, each* $x^* \in B^{-1}(0)$*, and each* $\beta > 0$*;*

(v) Suppose that $B^{-1}(0) \neq \emptyset$. *Then* $\langle x - J_\beta^B x, J_\beta^B x - w\rangle \geq 0$ *for each* $x \in H$*, each* $w \in B^{-1}(0)$*, and each* $\beta > 0$*.*

The next lemma gives a crucial characterization of the solution sets of the SVIP and the fixed point sets of the resolvent operator.

Lemma 2.4. [12] *Let* H_1 *and* H_2 *be real Hilbert spaces,* $A : H_1 \to H_2$ *be a bounded linear operator. Let* $\beta > 0$*,* $\gamma > 0$*,* $B_1 : H_1 \to 2^{H_1}$ *and* $B_2 : H_2 \to 2^{H_2}$ *be set-valued maximal monotone mappings. Given any* $x^* \in H_1$*.*

(i) If x^* *is a solution of (SVIP), then* $J_\beta^{B_1}(x^* - \gamma A^*(I - J_\beta^{B_2})Ax^*) = x^*$*.*

(ii) Suppose that $J_\beta^{B_1}(x^* - \gamma A^*(I - J_\beta^{B_2})Ax^*) = x^*$ *and the solution set of (SVIP) is nonempty. Then* x^* *is a solution of (SVIP).*

Lemma 2.5. [12] *Let* H_1 *and* H_2 *be real Hilbert spaces,* $A : H_1 \to H_2$ *be a bounded linear operator and* $\beta > 0$*. Let* $B : H_2 \to 2^{H_2}$ *be a set-valued maximal monotone mapping. Define a mapping* $T : H_1 \to H_1$ *by* $Tx := A^*(I - J_\beta^B)Ax$ *for each* $x \in H_1$*. Then*

(i) $\|(I - J_\beta^B)Ax - (I - J_\beta^B)Ay\|^2 \leq \langle Tx - Ty, x - y\rangle$ *for all* $x, y \in H_1$*;*

(ii) $\|A^*(I - J_\beta^B)Ax - A^*(I - J_\beta^B)Ay\|^2 \leq \|A\|^2 \cdot \langle Tx - Ty, x - y\rangle$ *for all* $x, y \in H_1$*.*

We also need the following tools in convergence analysis.

Lemma 2.6. [22] *Let* $\{a_n\}$*,* $\{b_n\}$ *and* $\{c_n\}$ *be real positive sequences such that*

$$a_{n+1} \leq (1 + c_n)a_n + b_n, \ n \geq 1.$$

If $\Sigma_{n=1}^{\infty} c_n < +\infty$ *and* $\Sigma_{n=1}^{\infty} b_n < +\infty$*, then* $\lim_{n\to+\infty} a_n$ *exists.*

Lemma 2.7. [16] *Let* $\{a_n\}$ *and* $\{\theta_n\}$ *be real positive sequences such that*

$$a_{n+1} \leq (1 + \theta_n)a_n + \theta_n a_{n-1}, \ n \geq 1.$$

Then, $a_{n+1} \leq K \cdot \prod_{i=1}^{n}(1 + 2\theta_i)$ *where* $K = \max\{a_1, a_2\}$*. Moreover, if* $\sum_{n=1}^{\infty} \theta_n < +\infty$*, then* $\{a_n\}$ *is bounded.*

Lemma 2.8. (Opial theorem [21]) *Let* C *be a nonempty subset of a real Hilbert space* H *and* $\{x_n\}$ *be a sequence in* H *that satisfies the following properties:*

(i) $\lim_{n\to\infty} \|x_n - x\|$ *exists for each* $x \in C$*;*

(ii) every sequential weak limit point of $\{x_n\}$ *is in* C*.*

Then $\{x_n\}$ *converges weakly to a point in* C*.*

3. Main Results

In this section, we introduce an inertial proximal algorithm using the self adaptive stepsize and prove the weak convergence theorem. We denote Ω is a the solution set of the SVIP and assume that Ω is nonempty. Let H_1 and H_2 be real Hilbert spaces, $A : H_1 \to H_2$ be a linear and bounded operator, and A^* be the adjoint operator of A. Let $B_1 : H_1 \to 2^{H_1}$ and $B_2 : H_2 \to 2^{H_2}$ be set-valued maximal monotone operators.

Algorithm 3.1. Let $\sigma_n \in (0,2)$, $\theta_1 > 0$, $0 < \varepsilon_n < 1$ and $\{\beta_n\}_{n\in\mathbb{N}}$ be a sequence in $(0,\infty)$. Given the iterates $\{x_{n-1}\}$ and $\{x_n\}$ for each $n \geq 1$.

Step 1 Compute the inertial step as follows:

$$w_n = x_n + \theta_n(x_n - x_{n-1}). \tag{3.1}$$

Step 2 Compute the proximal step as follows:

$$\begin{aligned} y_n &= J_{\beta_n}^{B_1}(w_n - \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n) \\ x_{n+1} &= y_n - \lambda_n A^*(I - J_{\beta_n}^{B_2})Ay_n \end{aligned} \tag{3.2}$$

where

$$\alpha_n = \frac{\sigma_n\|(I - J_{\beta_n}^{B_2})Aw_n\|^2}{\|A^*(I - J_{\beta_n}^{B_2})Aw_n\|^2 + \varepsilon_n} \quad \text{and} \quad \lambda_n = \frac{\sigma_n\|(I - J_{\beta_n}^{B_2})Ay_n\|^2}{\|A^*(I - J_{\beta_n}^{B_2})Ay_n\|^2 + \varepsilon_n}. \tag{3.3}$$

Theorem 3.1. *Let $\{x_n\}$ be a sequence generate by Algorithm 3.1. Assume that $\{\beta_n\}_{n\in\mathbb{N}}$ is a sequence in $[\beta,\infty)$ for some $\beta > 0$ and $\lim_{n\to\infty}\varepsilon_n = 0$. If $\sum_{n=1}^{\infty}\theta_n < \infty$, then $\{x_n\}$ weakly converges to a solution in Ω.*

Proof. Let $z \in \Omega$. Then $z \in B_1^{-1}(0)$ and $Az \in B_2^{-1}(0)$. Consider

$$\begin{aligned} \|x_{n+1} - z\|^2 &= \|y_n - \lambda_n A^*(I - J_{\beta_n}^{B_2})Ay_n - z\|^2 \\ &= \|y_n - z\|^2 - 2\lambda_n\langle y_n - z, A^*(I - J_{\beta_n}^{B_2})Ay_n\rangle \\ &\quad + \lambda_n^2\|A^*(I - J_{\beta_n}^{B_2})Ay_n\|^2. \end{aligned} \tag{3.4}$$

Since, $J_{\beta_n}^{B_2}Az = Az$, it follows that

$$\begin{aligned} \langle y_n - z, A^*(I - J_{\beta_n}^{B_2})Ay_n\rangle &= \langle y_n - z, A^*(I - J_{\beta_n}^{B_2})Ay_n - A^*(I - J_{\beta_n}^{B_2})Az\rangle \\ &= \langle Ay_n - Az, (I - J_{\beta_n}^{B_2})Ay_n - (I - J_{\beta_n}^{B_2})Az\rangle \\ &\geq \|(I - J_{\beta_n}^{B_2})Ay_n\|^2. \end{aligned} \tag{3.5}$$

Also, we have

$$\langle w_n - z, A^*(I - J_{\beta_n}^{B_2})Aw_n\rangle \geq \|(I - J_{\beta_n}^{B_2})Aw_n\|^2. \tag{3.6}$$

By Lemma 2.2 (iii), we get

$$\begin{aligned} \|y_n - z\|^2 &= \|J_{\beta_n}^{B_1}(w_n - \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n) - z\|^2 \\ &\leq \|w_n - \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n - z\|^2 - \|y_n - w_n + \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n\|^2 \\ &= \|w_n - z\|^2 - 2\alpha_n\langle w_n - z, A^*(I - J_{\beta_n}^{B_2})Aw_n\rangle + \alpha_n^2\|A^*(I - J_{\beta_n}^{B_2})Aw_n\|^2 \\ &\quad - \|y_n - w_n + \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n\|^2 \end{aligned} \tag{3.7}$$

Combine (3.5), (3.6), (3.7) into (3.4)

$$
\begin{aligned}
& \|x_{n+1}-z\|^2 \\
\leq\ & \|w_n-z\|^2-2\alpha_n\langle w_n-z, A^*(I-J_{\beta_n}^{B_2})Aw_n\rangle+\alpha_n^2\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\
& -\|y_n-w_n+\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2-2\lambda_n\langle y_n-z, A^*(I-J_{\beta_n}^{B_2})Ay_n\rangle \\
& +\lambda_n^2\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2 \\
\leq\ & \|w_n-z\|^2-2\alpha_n\|(I-J_{\beta_n}^{B_2})Aw_n\|^2+\alpha_n^2\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\
& -\|y_n-w_n+\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2-2\lambda_n\|(I-J_{\beta_n}^{B_2})Ay_n\|^2 \\
& +\lambda_n^2\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2 \\
=\ & \|w_n-z\|^2-\frac{2\sigma_n\|(I-J_{\beta_n}^{B_2})Aw_n\|^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n}\|(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\
& +\frac{\sigma_n^2(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{(\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n)^2}\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\
& -\|y_n-w_n+\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2-\frac{2\sigma_n\|(I-J_{\beta_n}^{B_2})Ay_n\|^2}{\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n}\|(I-J_{\beta_n}^{B_2})Ay_n\|^2 \\
& +\frac{\sigma_n^2(\|(I-J_{\beta_n}^{B_2})Ay_n\|^2)^2}{(\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n)^2}\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2 \\
\leq\ & \|w_n-z\|^2-\frac{2\sigma_n(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n}+\frac{\sigma_n^2(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n} \\
& -\|y_n-w_n+\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2-\frac{2\sigma_n(\|(I-J_{\beta_n}^{B_2})Ay_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n} \\
& +\frac{\sigma_n^2(\|(I-J_{\beta_n}^{B_2})Ay_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n} \\
=\ & \|w_n-z\|^2-(2-\sigma_n)\sigma_n\frac{(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n} \\
& -\|y_n-w_n+\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\
& -(2-\sigma_n)\sigma_n\frac{(\|(I-J_{\beta_n}^{B_2})Ay_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n}.
\end{aligned}
\tag{3.8}
$$

Since $\sigma_n\in(0,2)$, this implies that

$$\|x_{n+1}-z\|\leq\|w_n-z\|, \tag{3.9}$$

So, we obtain

$$
\begin{aligned}
\|x_{n+1}-z\| &\leq \|w_n-z\| \\
&= \|x_n+\theta_n(x_n-x_{n_1})-z\| \\
&\leq \|x_n-z\|+\theta_n\|x_n-x_{n-1}\| \\
&\leq \|x_n-z\|+\theta_n(\|x_n-z\|+\|x_{n-1}-z\|).
\end{aligned}
\tag{3.10}
$$

Hence, $\|x_{n+1}-z\| \leq (1+\theta_n)\|x_n - z\| + \theta_n\|x_{n-1}-z\|$. By Lemma 2.7, we obtain

$$\|x_{n+1}-z\| \leq K\prod_{i=1}^{n}(1+2\theta_i) \tag{3.11}$$

where $K=\max\{\|x_1-z\|, \|x_2-z\|\}$. By Lemma 2.7 and $\sum_{n=1}^{\infty}\theta_n < +\infty$, we obtain $\{x_n\}$ is bounded. So $\sum_{n=1}^{\infty}\theta_n\|x_n - x_{n-1}\| < +\infty$. By Lemma 2.6, we obtain $\lim_{n\to\infty}\|x_n - z\|$ exists.

Further, we have

$$\begin{aligned}\|w_n - z\|^2 &= \|x_n + \theta_n(x_n - x_{n-1}) - z\|^2 \\ &= \|x_n - z\|^2 + 2\theta_n\langle x_n - z, x_n - x_{n-1}\rangle + \theta_n^2\|x_n - x_{n-1}\|^2 \\ &\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \qquad (3.12)\end{aligned}$$

From (3.8) and (3.12), we have

$$\begin{aligned}\|x_{n+1}-z\|^2 &\leq \|x_n - z\|^2 + 2\theta_n\|x_n - z\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\ &\quad -(2-\sigma_n)\sigma_n\frac{(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n} \\ &\quad -\|y_n - w_n + \alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2 \\ &\quad -(2-\sigma_n)\sigma_n\frac{(\|(I-J_{\beta_n}^{B_2})Ay_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Ay_n\|^2+\varepsilon_n}. \qquad (3.13)\end{aligned}$$

From (3.13) and assumption of σ_n, it follows that

$$\lim_{n\to\infty}\frac{(\|(I-J_{\beta_n}^{B_2})Aw_n\|^2)^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n} = 0. \tag{3.14}$$

It is easy to check that $\{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|\}$ is bounded. Therefore, we obtain

$$\lim_{n\to\infty}\|(I-J_{\beta_n}^{B_2})Aw_n\| = 0, \tag{3.15}$$

Also, we get

$$\lim_{n\to\infty}\|(I-J_{\beta_n}^{B_2})Ay_n\| = 0. \tag{3.16}$$

From (3.15), we see that

$$\begin{aligned}\alpha_n\|A^*(I-J_{\beta_n}^{B_2})Aw_n\| &= \frac{\sigma_n\|(I-J_{\beta_n}^{B_2})Aw_n\|^2}{\|A^*(I-J_{\beta_n}^{B_2})Aw_n\|^2+\varepsilon_n}\|A^*(I-J_{\beta_n}^{B_2})Aw_n\| \\ &\to 0 \text{ as } n\to\infty. \qquad (3.17)\end{aligned}$$

From (3.13), we have

$$\lim_{n\to\infty}\|y_n - w_n + \alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\| = 0. \tag{3.18}$$

By (3.17) and (3.18), it implies that

$$\begin{aligned}\|y_n - w_n\| &= \|y_n - w_n + \alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n - \alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\| \\ &\leq \|y_n - w_n + \alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\| + \|\alpha_n A^*(I-J_{\beta_n}^{B_2})Aw_n\| \\ &\to 0 \text{ as } n\to\infty. \qquad (3.19)\end{aligned}$$

Hence, we have

$$\lim_{n\to\infty} \|y_n - w_n\| = 0. \tag{3.20}$$

From (3.1), we have

$$\lim_{n\to\infty} \|w_n - x_n\| = 0. \tag{3.21}$$

Also, from (3.3) and $\lim_{n\to\infty} \|(I - J_{\beta_n}^{B_2})Ay_n\| = 0$, we obtain

$$\lim_{n\to\infty} \|x_{n+1} - y_n\| = 0. \tag{3.22}$$

By Lemma 2.3 (i), (3.15) and (3.20), we have

$$\begin{aligned} \|Ay_n - J_{\beta_n}^{B_2} y_n\| &\leq \|Ay_n - J_{\beta_n}^{B_2} Ay_n - Aw_n + J_{\beta_n}^{B_2} w_n\| + \|Aw_n - J_{\beta_n}^{B_2} Aw_n\| \\ &\leq 2\|A\|\|y_n - w_n\| + \|Aw_n - J_{\beta_n}^{b_2} Aw_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.23}$$

By Lemma 2.3 (iii), we have

$$\begin{aligned} \|Ay_n - J_{\beta}^{B_2} Ay_n\| &\leq \|Ay_n - J_{\beta_n}^{B_2} Ay_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.24}$$

By Lemma 2.3 (i) and (3.15), we have

$$\begin{aligned} \|y_n - J_{\beta_n}^{B_1} w_n\| &= \|J_{\beta_n}^{B_1}(w_n - \alpha_n A^*(I - J_{\beta_n}^{B_2})Aw_n) - J_{\beta_n}^{B_1}(w_n)\| \\ &\leq \alpha_n \|A^*\|\|(I - J_{\beta_n}^{B_2})Aw_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.25}$$

From (3.20) and (3.25), we obtain

$$\begin{aligned} \|w_n - J_{\beta_n}^{B_1} w_n\| &= \|w_n - y_n + y_n - J_{\beta_n}^{B_1} w_n\| \\ &\leq \|w_n - y_n\| + \|y_n - J_{\beta_n}^{B_1} w_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.26}$$

Also, by Lemma 2.3 (iii), we get

$$\begin{aligned} \|w_n - J_{\beta}^{B_1} w_n\| &\leq \|w_n - J_{\beta_n}^{B_1} w_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.27}$$

From (3.15) and (3.21), we obtain

$$\begin{aligned} \|Ax_n - J_{\beta}^{B_2} Ax_n\| &= \|Ax_n - J_{\beta}^{B_2} Ax_n - Aw_n + J_{\beta}^{B_2} Aw_n\| + \|Aw_n - J_{\beta}^{B_2} Aw_n\| \\ &\leq 2\|A\|\|x_n - w_n\| + \|Aw_n - J_{\beta}^{B_2} Aw_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.28}$$

From (3.21) and (3.27), we obtain

$$\begin{aligned} \|x_n - J_{\beta}^{B_1} x_n\| &= \|x_n - w_n + w_n - J_{\beta}^{B_1} w_n + J_{\beta}^{B_1} w_n - J_{\beta}^{B_1} x_n\| \\ &\leq \|x_n - w_n\| + \|w_n - J_{\beta}^{B_1} w_n\| + \|J_{\beta}^{B_1} w_n - J_{\beta}^{B_1} x_n\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{3.29}$$

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ and $x^* \in H_1$ such that $x_{n_k} \rightharpoonup x^*$. Since A is a bounded linear operator, we have $Ax_{n_k} \rightharpoonup Ax^*$. By (3.28), (3.29), Lemma 2.1 and Lemma 2.3 (ii), we have $x^* \in \Omega$. By Lemma 2.8, we conclude that the sequence $\{x_n\}$ converges weakly to a point in Ω. This completes the proof. ∎

4. Application

In this section, we apply the proposed algorithm, Algorithm 3.1 to the split feasibility problem (SFP) which is the problem of finding $x^* \in H_1$ such that

$$x^* \in C \text{ and } Ax^* \in Q,$$

where H_1 and H_2 are real Hilbert spaces, C and Q are nonempty closed convex subsets of H_1 and H_2, respectively, and $A : H_1 \to H_2$ is a linear and bounded operator with adjoint operator A^*.

Let H be a Hilbert space and let $g : H \to (-\infty, \infty]$ be a proper, lower semicontinuous and convex function. The subdifferential ∂g of g is defined by

$$\partial g(x) = \{z \in H : g(x) + \langle z, y - x\rangle \leq g(y), \forall y \in H\} \tag{4.1}$$

for all $x \in H$. Let C be a nonempty closed convex subset of H, and $\imath_C$ be the indicator function of C defined by

$$\imath_C x = \begin{cases} 0 & x \in C, \\ \infty & x \notin C. \end{cases} \tag{4.2}$$

The normal cone $N_C u$ of C at u is defined by

$$N_C u = \{z \in H : \langle z, v - u\rangle \leq 0, \forall v \in C\}. \tag{4.3}$$

Then, $\imath_C$ is a proper, lower semicontinuous and convex function on H. See [3, 27]. Moreover, the subdifferential $\partial \imath_C$ of $\imath_C$ is a maximal monotone mapping. In this connection, we can define the resolvent $J_\beta^{\partial \imath_C}$ of $\partial \imath_C$ for $\beta > 0$ by

$$J_\beta^{\partial \imath_C} x = (I + \beta \partial \imath_C)^{-1} x \tag{4.4}$$

for all $x \in H$. Hence, we see that

$$\begin{aligned} \partial \imath_C x &= \{z \in H : \imath_C x + \langle z, y - x\rangle \leq \imath_C y, \forall y \in H\} \\ &= \{z \in H : \langle z, y - x\rangle \leq 0, \forall y \in C\} \\ &= N_C x \end{aligned} \tag{4.5}$$

for all $x \in C$. Hence, for each $\beta > 0$, we obtain the following relation:

$$\begin{aligned} u = J_\beta^{\partial \imath_C} &\Leftrightarrow x \in u + \beta \partial \imath_C u \\ &\Leftrightarrow x - u \in \beta N_C u \\ &\Leftrightarrow \langle x - u, y - u\rangle \leq 0, \forall y \in C \\ &\Leftrightarrow u = P_C x. \end{aligned} \tag{4.6}$$

So we consequently obtain the following results.

Algorithm 4.1. Let $\sigma_n \in (0, 2)$, $\theta_1 > 0$, $0 < \varepsilon_n < 1$. Given the iterates x_{n-1} and x_n for each $n \geq 1$.

Step 1 Compute the inertial step as follows:

$$w_n = x_n + \theta_n (x_n - x_{n-1}). \tag{4.7}$$

Step 2 Compute the P_C step as follows:

$$\begin{aligned} y_n &= P_C(w_n - \alpha_n A^*(I - P_Q)Aw_n) \\ x_{n+1} &= y_n - \lambda_n A^*(I - P_Q)Ay_n \end{aligned} \tag{4.8}$$

where

$$\alpha_n = \frac{\sigma_n \|(I - P_Q)Aw_n\|^2}{\|A^*(I - P_Q)Aw_n\|^2 + \varepsilon_n} \quad \text{and} \quad \lambda_n = \frac{\sigma_n \|(I - P_Q)Ay_n\|^2}{\|A^*(I - P_Q)Ay_n\|^2 + \varepsilon_n}. \tag{4.9}$$

Theorem 4.1. *Let $\{x_n\}$ be a sequence generate by Algorithm 4.1. Assume that* $\lim_{n\to\infty} \varepsilon_n = 0$. *If $\sum_{n=1}^{\infty} \theta_n < \infty$, then $\{x_n\}$ weakly converges to a solution in Ω.*

We next consider the example by using Algorithm 4.1 to data classification problems, which based on a learning technique called extreme learning machine (ELM).

Let $\{(x_n, y_n) : x_n \in \mathbb{R}^N, y_n \in \mathbb{R}^M, n = 1, 2, 3, ..., K\}$ be a training set of K distinct samples, x_n is an input training data and y_n is a training target. For the output of ELM with single hidden layer at the i-th hidden node is

$$h_i(x) = U(a_i \cdot x + b_i),$$

where U is an activation function, a_i is the weight at the i-th hidden node and b_i is the bias at the i-th hidden node.

The output function with L hidden nodes is the single-hidden layer feed forward neural networks (SLFNs)

$$O_n = \sum_{i=1}^{L} \omega_i h_i(x_n),$$

where ω_i is the optimal output weight at the i-th hidden node. The hidden layer output matrix A is defined by

$$A = \begin{bmatrix} U(a_1 \cdot x_1 + b_1) & \cdots & U(a_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ U(a_1 \cdot x_K + b_1) & \cdots & U(a_L \cdot x_K + b_L) \end{bmatrix}.$$

The main aim of ELM is to calculate an optimal weight $\omega = [\omega_1, ..., \omega_L]^T$ such that $A\omega = \chi$, where $\chi = [t_1, ..., t_K]^T$ is the training target data. We find the solution ω for avoiding overfitting of our model via constrained optimization problem. It can be formulated as follows:

$$\min_{\omega \in \mathbb{R}^L} \|A\omega - \chi\|_2^2 \text{ such that } \|\omega\|_1 \le \tau, \tag{4.10}$$

where $\tau > 0$ is a given constant. In particular, if $C = \{\omega \in \mathbb{R}^L : \|\omega\|_1 \le \tau\}$ and $Q = \{\chi\}$, then (4.10) can be consider as the SFP.

We use the sigmoid as the activation function and the hidden nodes $L = 200$. Accuracy, Recall and Precision. A measure of the algorithm can correctly predict cases into their correct category is measured by the classifier accuracy, the percentage of a certain class correctly identified is presented by recall (also known as sensitivity), and the percentage of quality of a positive prediction made by the algorithm is measured by precision. The formulation of three measures [29] is defined as follow:

$$\text{Precision (Pre)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%$$

$$\text{Recall (Rec)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

$$\text{Accuracy (Acc)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \times 100\%$$

where a confusion matrix for original and predicted classes are shown in terms of TP: = True Positive which mean accurately classified, TN: =True Negative shows inaccurately classified, FP: = False Positive.

For the loss of an example, it is computed by the multi-class cross entropy loss function:

$$\text{Loss} = -\sum_{i=1}^{K} y^k \log \hat{y}^k,$$

where y^k is 0 or 1, indicating whether class label k is the correct classification and $\hat{y}^k$ is a probability of class y^k and K is the number of scalar values in the model output.

The regularization parameter is $\tau = 10^{-5}$ and $\sigma = 1.2$. The stopping criteria is the multi-class cross entropy (Loss=0.12). We consider five cases for the different parameters θ_n as follows:

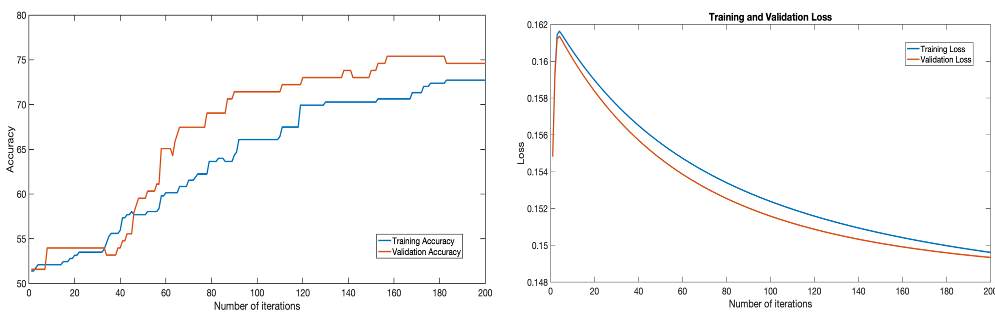
Case 1: $\theta_n = 0$, Case 2: $\theta_n = \dfrac{1}{n^2}$, Case 3: $\theta_n = \dfrac{1}{n^2+1}$, Case 4: $\theta_n = \dfrac{1}{\|x_n - x_{n-1}\|^3 + n^3}$, Case 5: $\theta_n = \dfrac{10^{10}}{\|x_n - x_{n-1}\|^3 + n^3 + 10^{10}}$. Then the result as follows Table 1.

TABLE 1. Numerical results for different parameter θ_n

θ_n	Iter	Time	Pre(%)	Rec(%)	Acc(%)
Case 1	238	0.5067	52.38	52.38	76.1905
Case 2	238	0.4643	52.38	52.38	76.1905
Case 3	18	0.1524	55.56	55.56	77.7778
Case 4	238	0.5205	52.38	52.38	76.1905
Case 5	19	0.0446	55.56	55.56	77.7778

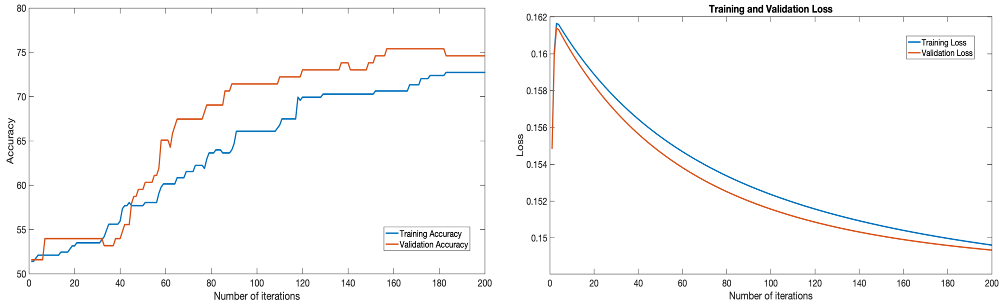
From Table 1, we see that Case 3 ($\theta_n = \dfrac{1}{n^2+1}$) has number of iterations less than other cases in the proposed algorithm 3.1 at performance of accuracy 77.7778%.

Next, we show graphs of the accuracy and loss of training data and testing data for overfitting of Algorithm 3.1.



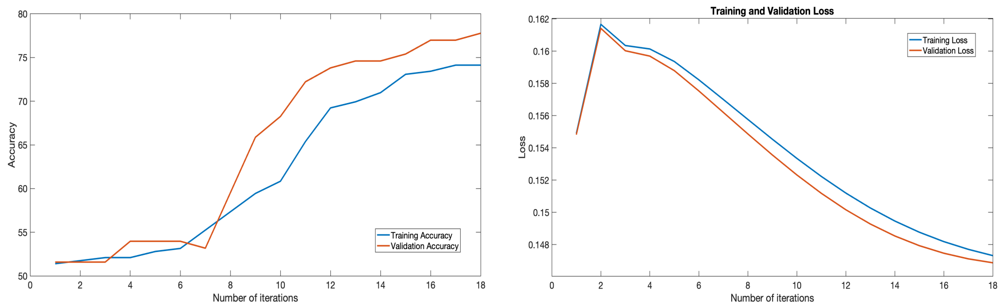
(A) Acc (B) Loss

FIGURE 1. Graph accuracy and loss of Algorithm 3.1 in Case 1



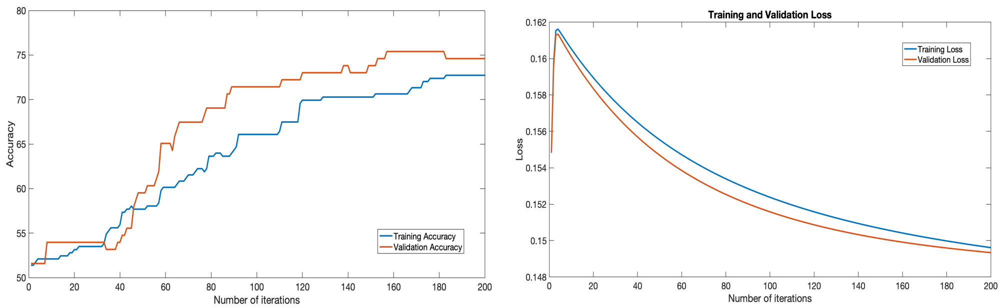
(A) Acc (B) Loss

FIGURE 2. Graph accuracy and loss of Algorithm 3.1 in Case 2



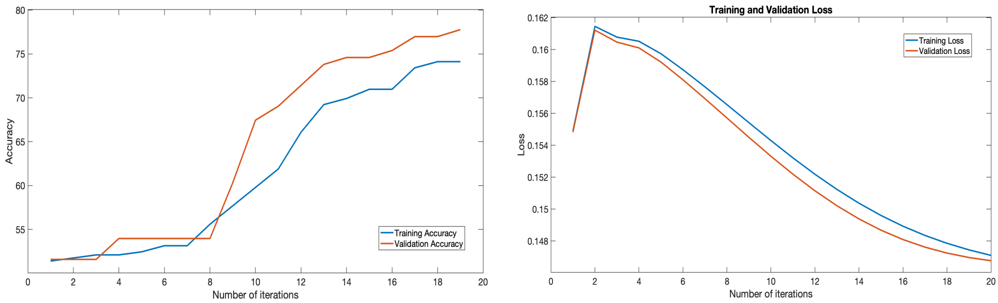
(A) Acc (B) Loss

FIGURE 3. Graph accuracy and loss of Algorithm 3.1 in Case 3



(A) Acc (B) Loss

FIGURE 4. Graph accuracy and loss of Algorithm 3.1 in Case 4



(A) Acc (B) Loss

FIGURE 5. Graph accuracy and loss of Algorithm 3.1 in Case 5

From Figures 1-5, we observe that our Algorithm 3.1 has good fitting model show that our algorithm suitably learns the training dataset and generalizes well to a hold-out dataset.

5. Conclusions and Discussion

In this paper, we introduced an iterative method with inertial term using the self adaptive stepsize. We also proved weak convergence theorem under some suitable conditions for solving the split variational inclusion problem. We also presented a numerical experiment in data the classification problem and provided a result of different parameters showing our algorithm has efficiency. The application of this study was to classify the TPACK level of pre-service mathematics teachers through self-assessment-based measures. The results indicate that the proposed method attained an accuracy of categorization of 77.7778%. Accordingly, GPA, gender, CK, PK, TK, PCK, TCK, and TPK are significant predictors of their TPACK level.

Reviewing the results from this study to those of other studies on the TPACK level of pre-service mathematics teachers revealed that gender is one of the most accurate predictors of teachers' intentions to incorporate technology in their classrooms (Anderson,Arumugam). Undoubtedly, Content Knowledge (CK) and Pedagogical Knowledge (PK) directly predict TPACK level, as well as GPA, which is a reflection of CK and PK.

Moreover, broad technological abilities and knowledge, such as Technological Knowledge (TK), effectively predict TPACK levels (Bonafini,Niess,Niess2). In addition, the integration of technology knowledge and other knowledge as Technological Content Knowledge (TCK) and Technological Pedagogical Knowledge (TPK), are unquestionable predictors of TPACK (Niess2,Bonafini,Mouza).

Based on these findings, it is possible to predict future TPACK level using this method. By projecting the TPACK level of pre-service mathematics teachers, teacher educators can assess and improve their working approaches and expertise. Given that there are around four years between teacher education programs, it is simpler to appreciate the proposed strategy's implications.

The practical contribution of this research is a curriculum revision policy for mathematics teacher education programs. Specifically, the program should offer more TCK and

TPK courses, that means the revised curriculum should place a stronger focus on technological expertise in mathematics, and in teaching. In addition, the finding describes the concept of the required curricular pattern in terms of balancing courses in pedagogical knowledge, content knowledge, and technology knowledge.

The results indicate that machine learning approaches can be utilized to forecast the teacher knowledge level of pre-service mathematics teachers. This research can help teacher educators identify pre-service mathematics teachers with below-average or above-average TPACK. Moreover, such data-driven research are crucial for developing a projected teacher knowledge analysis framework in teacher education and for contributing to the creation of policies.

Acknowledgements

We would like to thank Lampang Rajabhat University for supporting the raw dataset to be analyzed in our main result of this paper. Watcharaporn Cholamjiak would like to thank National Research Council of Thailand and University of Phayao (N42A650334); Thailand Science Research and Innovation, University of Phayao (FF66-UoE). Nipa Jun-on would like to thank the revenue budget in 2023, Lampang Rajabhat University.

References

[1] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, SIAM J. Optim. 14 (2004) 773–782.

[2] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (1) (2001) 3–11.

[3] R.P. Agarwal, D. O'Regan, D.R. Sahu, Fixed Point Theory for Lipschitzian-Type Mapping with Applications, Springer: New York, 2009.

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

[5] R.I. Bot, E.R. Csetnek, An inertial alternating directionmethod of multipliers, Minimax Theory Appl. 1 (2016) 29–49.

[6] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, Inverse Probl. 18 (2) (2002) 441–453.

[7] C. Byrne, A unified treatment of some iterative algorithms in signal processing and image reconstruction, Inverse Probl. 20 (1) (2004) 103–120.

[8] C. Byrne, Y. Censor, A. Gibali, Weak and strong convergence of algorithms for the split common null point problem, J. Nonlinear Convex Anal. 13 (2011) 759–775.

[9] Y. Censor, A. Gibali, S. Reich, Algorithms for the split variational inequality problem, Numer. Algor. 59 (2) (2012) 301–323.

[10] Y. Censor, T. Bortfeld, B. Martin, A unified approach for inversion problems in intensity-modulated radiation therapy, Phys. Med. Biol. 51 (10) (2006) 2353–2365.

[11] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a product space, Numer. Algor. 8 (2) (1994) 221–239.

[12] C.S. Chuang, Strong convergence theorems for the split variational inclusion problem in Hilbert spaces, Fixed Point Theory Appl. 350 (2013).

[13] C.S. Chuang, Hybrid inertial proximal algorithm for the split variational inclusion problem in Hilbert spaces with applications, Optimization 66 (5) (2017) 777–792.

[14] Y. Dang, Y. Gao, The strong convergence of a KM-CQ-like algorithm for a split feasibility problem, Inverse Problems 27 (1) (2010) 015007.

[15] K. Goebel, W.A. Kirk, Topics in Metric Fixed Point Theory, Cambridge: Cambridge University Press, 1990.

[16] A. Hanjing, S. Suantai, A fast image restoration algorithm based on a fixed point and optimization method, Mathematics 8 (3) (2020) 378.

[17] L.J. Lin, Y.D. Chen, C.S. Chuang, Solutions for a variational inclusion problem with applications to multiple sets split feasibility problems, Fixed Point Theory Appl. 2013 (2013) Article no. 333.

[18] A. Moudafi, Split monotone variational inclusions, J. Optim Theory Appl. 150 (2) (2011) 275–283.

[19] G. Marino, H.K. Xu, Convergence of generalized proximal point algorithms, Commun. Pure Appl. Anal. 3 (4) (2004) 791–808.

[20] T.L. Nguyen, Y. Shin, Deterministic sensing matrices in compressive sensing: A survey, The Scientific World Journal (2013).

[21] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, Bull. Am. Math. Soc. 73 (4) (1967) 591–597.

[22] M.O. Osilike, S.C. Aniagbosor, B.G. Akuchu, Fixed points of asymptotically demicontractive mappings in arbitrary Banach spaces, Panam. Math. J. 12 (2) (2002) 77–88.

[23] N. Pakkaranang, P. Kumam, Y.I. Suleiman, B. Ali, Bounded perturbation resilience of viscosity proximal algorithm for solving split variational inclusion problems with applications to compressed sensing and image recovery, Mathematical Methods in the Applied Sciences 45 (8) (2022) 4085–4107.

[24] B.T. Polyak, Some methods of speeding up the convergence of iteration methods, USSR Comput. Math. and Math. Phy. 4 (5) (1964) 1–17.

[25] L. Shulman, Those who understand: Knowledge growth in teaching, Educ. Res. 15 (2) (1986) 4–14.

[26] L. Shulman, Knowledge and teaching: Foundations of the new reform, Harv. Educ. Rev. 57 (1987) 1–22.

[27] W. Takahashi, Introduction to Nonlinear and Convex Analysis, Yokohama Tokohama Publishers, Inc. Yokohama, 2009.

[28] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B. Stat. Methodol. 58 (1996) 267–288.

[29] T. Thomas, N. Pradhan, V.S. Dhaka, Comparative analysis to predict breast cancer using machine learning algorithms: A survey. 2020 international conference on inventive computation technologies (ICICT), IEEE, February (2020).

[30] N. Wairojjana, N. Pholasa, N. Pakkaranang, On strong convergence theorems for a viscosity-type tseng's extragradient methods solving quasimonotone variational inequalities, Nonlinear Functional Analysis and Applications 27 (2) (2022) 381–403.

[31] H.K. Xu, Iterative methods for the split feasibility problem in infinite-dimensional Hilbert spaces, Inverse Prob. 26 (2010) 105018.