An inertial subgradient extragradient method for approximating a solution of an equilibrium problem in an Hadamard manifold

Simeon Reich and Olawale. K. Oyewole

sreich@technion.ac.il
oyewoleok@campus.technion.ac.il

A workshop on Nonlinear Functional Analysis and Its Applications in memory of Professor Ronald E. Bruck.
Technion-Israel institute of technology, Haifa, Israel.

6th April 2022

The term Equilibrium problem by Blum and Oetlli [5], 1994.

Further work on the EP given by Muu and Oetlli [27].

Standard examples of the EP.
- Minimization problem.
- Variational inequality.
- Fixed point problem.
Let $g : K \times K \to \mathbb{R}$ be a bifunction such that $g(x, x) = 0$ for all $x \in K$, where $K$ is a nonempty subset of a topological space $X$. Then, the EP is to find a point $x \in K$ such that

$$g(x, y) \geq 0 \quad \forall y \in K. \quad (1)$$

We denote by $\text{Sol}(g, K)$, the solution set of the EP (1).
Why Hadamard Manifold?

The study of variational inequality, equilibrium and other related optimization problems are receiving several attentions of researchers in the framework of Riemannian manifolds. Thus, methods and ideas are being extended from linear settings to this more generalized settings.
So why Hadamard manifold?

- convexity
- constraints

Consistency of the problem (using Fan’s KKM lemma).

By replacing $M$ with $X$ in (1), we arrive at Colao et al. [12] formulation of the EP.

After the work of Colao et al. [12], there has been

- (2016) Tang et al. [38],
- (2017) Salahuddin [34],
- (2016) Zhou and Huang [41].
Iterative methods

The development of effective iterative algorithm for approximating the solution of an optimization problem is another interesting direction. Some iterative algorithms for EP

- (1976) Extragradien Method (EGM) by Korpelevich [22]
- Tseng extragradien method
- Projection extragradien method and so on.

For EPs, Quoc et al. [39] (2008) introduced an extragradien-like for approximating a solution of a pseudomonotone EP.
Recent works in this direction

1. (2014) Nguyen et al. [28] introduced a method for finding a common solution of a fixed point and equilibrium problem based on the extragradient method in Quoc et al. [39]

2. (2021) Habib et al. [31] introduced an inertial viscosity subgradient extragradient algorithm for solving the equilibrium problem.
(2016) Cruz et al. [14] extended the result of Nguyen et al. [28] to the settings of Hadamard manifold by considering the following algorithm: Let $x_1 \in K$ and $\lambda_n > 0$, compute

$$
\begin{align*}
    y_n &= \arg \min_{y \in M} \left\{ g(x_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\}, \\
    x_{n+1} &= \arg \min_{y \in M} \left\{ g(y_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\},
\end{align*}
$$

such that $0 \leq \lambda_n < \mu < \min\left\{ \frac{1}{c_1}, \frac{1}{c_2} \right\}$ where $c_1 > 0$ and $c_2 > 0$ are the Lipschitz constants with respect to the bifunction $g$. By replacing $d(x, y)$ with $\|x - y\|$ we obtain the method of Nguyen et al. [28].
Fan et al. [17] method.  
Choose $x_1 \in K$ and $\lambda > 0$, $\mu \in (0, 1)$. Given the current iterate $x_n \in K$ and $\lambda_n (n \geq 0)$. Calculate $x_{n+1}$, $\lambda_{n+1}$ as

$$
\begin{align*}
    y_n &= \arg\min_{y \in K} \left\{ f(x_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\}, \\
    x_{n+1} &= \arg\min_{y \in K} \left\{ f(y_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\}
\end{align*}
$$

and

$$
\lambda_{n+1} = \left\{ \lambda_n, \frac{\mu [d^2(x_n, y_n) + d^2(x_{n+1}, y_n)]}{2[g(x_n, x_{n+1}) - g(x_n, y_n) - g(y_n, x_{n+1})]^+} \right\}.
$$


Results in Manifold

Choose an initial point $x_0 \in K$ and $\lambda$ such that $0 < \lambda < \min \left\{ \frac{1}{2c_1}, \frac{1}{2c_2} \right\}$. Assume that $x_n \in K$ and we calculate $x_{n+1} \in K$ as follows:

$$
\begin{align*}
y_n &= \arg \min_{y \in K} \left\{ f(x_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\}, \\
x_{n+1} &= \arg \min_{y \in T_n} \left\{ f(y_n, y) + \frac{1}{2\lambda_n} d^2(x_n, y) \right\}
\end{align*}
$$

where $T_n = \{ y \in M : \langle \exp^{-1}_{y_n} x_n - \lambda v_n, \exp^{-1}_{y_n} y \rangle \leq 0 \}$ and $v_n \in \partial_2 g(x_n, y_n)$. 
Methods:

(1) Subgradient method of Censor et al. [9].
(2) Viscosity approach [2, 15, 20].
(3) Colao et al. [12], Ali-Akbari [1], Fan et al. [17].
(4) Inertial technique [29].

Applications:

1. Competitive exchange economy
2. Product pricing
3. Fractional programming and so on.
Let $M$ be a $m$-dimensional manifold and $x \in M$, let $T_x M$ be the tangent space of $M$ at $x \in M$. We denote by $TM = \bigcup_{x \in M} T_x M$ the tangent bundle of $M$. An inner product $\mathcal{R}\langle \cdot, \cdot \rangle$ is called the Riemannian metric on $T_x M$. The corresponding norm to the inner product $\mathcal{R}_x \langle \cdot, \cdot \rangle$ on $T_x M$ is denoted by $\| \cdot \|_x$.

Given a piecewise smooth curve $\gamma : [a, b] \rightarrow M$ joining $x$ to $y$ (i.e $\gamma(a) = x$ and $\gamma(b) = y$), we define the length $l(\gamma)$ of $\gamma$ by

$$l(\gamma) = \int_a^b \| \gamma'(t) \| dt.$$ 

Let $\gamma$ be a smooth curve in $M$. A vector field $X$ along $\gamma$ is said to be parallel if $\nabla_{\gamma'} X = 0$, where 0 is the zero tangent vector (see [33]).
Definitions and Lemmas

1. [33] For a complete Riemannian manifold $M$, then the exponential map $\exp_x : T_x M \to M$ at $x \in M$ is defined by $\exp_x v = \gamma_v(1, x)$, $\forall v \in T_x M$, where $\gamma_v(\cdot, x)$ is the geodesic starting from $x$ with velocity $v$ (i.e. $\gamma_v(0, x) = x$ and $\gamma'_v(0, x) = v$). Then, for any $t$, we have $\exp_x tv = \gamma_v(t, x)$ and $\exp_x 0 = \gamma_v(0, x) = x$.

2. The mapping $\exp_x$ is differentiable on $T_x M$ for every $x \in M$. The exponential map has an inverse $\exp_x^{-1} : M \to T_x M$. For any $x, y \in M$, we have $d(x, y) = \| \exp_x^{-1} x \| = \| \exp_y^{-1} y \|$.

3. A complete simply connected Riemannian manifold of nonpositive sectional curvature is said to be an Hadamard manifold. From now, we denote by $M$ a finite dimensional Hadamard manifold.
Definitions and Lemmas

We now present some useful results and definition which will be useful in the convergence analysis of our main result.

**Proposition**

[33]. Let \( x \in M \). The exponential mapping \( \exp_x : T_x M \rightarrow M \) is a diffeomorphism, for any two points \( x, y \in M \), there exists a unique normalized geodesic joining \( x \) to \( y \), which is expressed by the formula

\[
\gamma(t) = \exp_x t \exp_x^{-1} y, \quad \forall \ t \in [0, 1].
\]
Definitions and Lemmas

The following definitions can be found in [5, 25]. Let $K$ be a nonempty convex subset of $M$. A bifunction $g : M \times M \to \mathbb{R}$ is said to be

(i) monotone on $K$, if

$$\langle g(x, y) + g(y, x) \rangle \leq 0, \quad \forall x, y \in K;$$

(ii) pseudomontone on $K$, if

$$g(x, y) \geq 0 \Rightarrow g(y, x) \leq 0, \quad \forall x, y \in K;$$

(iii) Lipschitz-type continuous, if there exist constants $c_1 > 0$ and $c_2 > 0$, such that

$$g(x, y) + g(y, z) \geq g(x, z) - c_1 d^2(x, y) - c_2 d^2(y, z) \quad \forall x, y, z \in K.$$
Definitions and Lemmas

For solving the EP (1), we make the following assumptions for $g$ on $K$:

(A1) $g$ is pseudomonotone on $K$ and $g(x, x) = 0$ for all $x \in M$;

(A2) $g(\cdot, y)$ is upper semicontinuous for all $y \in M$;

(A3) $g(x, \cdot)$ is convex and subdifferentiable for all fixed $x \in M$;

(A4) $g$ satisfies the Lipschitz-type condition on $M$ i.e.

$$g(x, y) + g(y, z) \geq g(x, z) - c_1 d^2(x, y) - c_2 d^2(y, z).$$
Definitions and Lemmas

**Proposition**

Let $M$ be an Hadamard manifold and $x \in M$. Let $\rho_x(y) = \frac{1}{2}d^2(x,y)$, then $\rho_x(y)$ is strictly convex and its gradient at $y$ is given by

$$\partial \rho_x(y) = -\exp_y^{-1} x.$$ 

**Proposition**

Let $K$ be a nonempty convex subset of an Hadamard manifold $M$ and $h : K \to \mathbb{R}$ be a convex subdifferentiable and lower semicontinuous on $K$. Then a point $x$ solves the convex minimization problem

$$\min_{x \in K} h(x) \iff 0 \in \partial h(x) + N_K(x).$$
[37] Let $x, y \in K$ and $\lambda \in [0, 1]$. Then, the following properties hold on $K$.

(i) $\|\lambda x + (1 - \lambda)y\|^2 = \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \lambda(1 - \lambda)\|x - y\|^2$;

(ii) $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y \rangle + \|y\|^2$;

(iii) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle$. 
Lemma

[32] Let \( \{a_n\} \) be a sequence of nonnegative real numbers, \( \{\alpha_n\} \) be a sequence of real numbers in \((0, 1)\) such that \( \sum_{n=1}^{\infty} \alpha_n = \infty \) and \( \{b_n\} \) be a sequence of real numbers. Assume that

\[
a_{n+1} \leq (1 - \alpha_n)a_n + \alpha_nb_n, \quad \forall \ n \geq 1.
\]

If \( \lim \sup_{k \to \infty} b_{n_k} \leq 0 \) for every subsequence \( \{a_{n_k}\} \) of \( \{a_n\} \) satisfying the condition

\[
\lim \inf_{k \to \infty} (a_{n_k+1} - a_{n_k}) \geq 0,
\]

then \( \lim_{n \to \infty} a_n = 0. \)
In this section, we propose a strong convergent algorithm for approximating a solution of the EP (1) and then discuss its convergence analysis. The solution set $\text{Sol}(g, K)$ is closed and convex [12, 34]. We assume that $\text{Sol}(g, K)$ is nonempty.
Algorithm

Inertial subgradient extragradient method for EP(ISEMIP)

Initialization: Choose $x_0, x_1 \in K$, $\lambda_1 > 0$, $\mu \in (0, 1)$, a nonnegative sequence of real numbers $\{\alpha_n\}$ such that and
\[
\sum_{n=1}^{\infty} \alpha_n < +\infty \quad \text{and} \quad \beta_n \subset (0, 1)
\]
satisfying

(C1) $\lim_{n \to \infty} \beta_n = 0$ and $\sum_{n=1}^{\infty} \beta_n = \infty$.

Step 1: Given $x_n, x_{n-1}$ and $\lambda_n$. Compute

\[
\begin{align*}
w_n &= \gamma_n^0(\theta_n), \\
y_n &= \arg\min_{y \in M} \left\{ g(w_n, y) + \frac{1}{2\lambda_n}d^2(w_n, y) \right\}.
\end{align*}
\]
Algorithm

If \( y_n = w_n \), then stop. Otherwise go to the next step.

**Step 2:** Define the half space \( T_n \) by

\[
T_n := \{ y \in M : \langle \exp_{y_n}^{-1} w_n - \lambda_n v_n, \exp_{y_n}^y \rangle \leq 0 \}\]

with \( v_n \in \partial_2 g(w_n, y_n) \) and compute

\[
z_n = \arg \min_{y \in T_n} \left\{ g(y_n, y) + \frac{1}{2\lambda_n} d^2(w_n, y) \right\}.
\] (3)

**Step 3:** Compute

\[
x_{n+1} = \gamma_n^1 (1 - \beta_n), \quad \forall \ n \geq 0,
\] (4)
Main result

Algorithm

where $\gamma^1_n : [0, 1] \to M$ is the geodesic joining $f(x_n)$ to $z_n$, that is $\gamma^1_n(0) = f(x_n)$ and $\gamma^1_n(1) = z_n$ for all $n \geq 0$. Let $d_n = g(w_n, z_n) - g(w_n, y_n) - g(y_n, z_n)$. Then,

$$
\lambda_{n+1} = \begin{cases} 
\min \left\{ \lambda_n + \alpha_n, \frac{\mu [d^2(y_n, w_n) + d^2(z_n, y_n)]}{2d_n} \right\}, & d_n > 0, \\
\lambda_n + \alpha_n, & \text{otherwise.}
\end{cases}
$$

Set $n := n + 1$ and return to Step 1.
Remark

We observe from Algorithm 6, that the method is self adaptive with the step-size allowed to increase from iteration to iteration unlike the monotone decreasing sequence step-size in [31]. Thus, the dependence of the bifunction $g$ on the Lipschitz constants is dispensed with.
Main result

The set $K$ is a subset of $T_n$. This claim is easily obtained by the Proposition 3, the Normal cone and subdifferential definition.

Lemma

The sequence $\{x_n\}$ defined recursively by Algorithm 6 satisfies the inequality

$$d^2(z_n, p) \leq d^2(w_n, p) - \left(1 - \frac{\mu \lambda_n}{\lambda_{n+1}}\right) \left[d^2(w_n, y_n) + d^2(y_n, z_n)\right].$$
In the next result, we show that the sequence \( \{x_n\} \) generated by Algorithm 6 has a weak limit.

**Lemma**

Let \( f : K \rightarrow K \) be a \( \kappa \)-contraction mapping. The sequence \( \{x_n\} \) generated by Algorithm 6 is bounded.

We present our main theorem result in the next slide with some highlight of the proof given.
Suppose condition (A1)-(A4) and let $f : K \to K$ be a $\kappa$-contraction, then the sequence $\{x_n\}$ generated by Algorithm 6 converges strongly to a point $p = P_{\text{Sol}(g,K)} f(p) \in \text{Sol}(g,K)$, where $P_{\text{Sol}(g,K)}$ is the projection of $K$ onto $\text{Sol}(g,K)$. 

S. Reich and O.K. Oyewole
Application to variational inequalities

Suppose

\[
g(x, y) = \begin{cases} 
\langle Gx, \exp_{\frac{1}{x}} y \rangle, & \text{if } x, y \in K, \\
+\infty, & \text{otherwise,}
\end{cases}
\]

where \( G : K \to M \) is a mapping, then the equilibrium problem (1) reduces to the variational inequality (VIP):

\[
\text{Find } x \in K \text{ such that } \langle Gx, \exp_{\frac{1}{x}} y \rangle \geq 0, \quad \forall y \in K. \quad (6)
\]

We denote the set of solution of VIP (6) by \( VIP(G, K) \).
The mapping $G : K \rightarrow M$ is said to be pseudomonotone if

$$\langle Gx, y - x \rangle \geq 0 \Rightarrow \langle Gy, y - x \rangle \geq 0, \ x, y \in K.$$ 

Assume that the function $G$ satisfies the following conditions:

(V1) The function $G$ is pseudomonotone on $K$ with $\text{VIP}(G, K) \neq \emptyset$.

(V2) $G$ is $L$-Lipschitz continuous, that is

$$\|Py, x Gx - Gy\| \leq \|x - y\|, \ x, y \in K,$$

where $Py, x$ is a parallel transport (see [21]).

(V3) $\limsup_{n \to \infty} \langle Gx_n, \exp^{-1}_{x_n} y \rangle \leq \langle Gp, \exp^{-1}_{p} y \rangle$ for every $y \in K$ and $\{x_n\} \subset K$ such that $x_n \rightharpoonup p$. 

S. Reich and O.K. Oyewole
Application formulation

By replacing the proximal term $\arg \min_{y \in M} \left\{ g(x, y) + \frac{1}{2\lambda_n} d^2(x, y) \right\}$ with $P_K(\exp_x(-\lambda_n G(x)))$, where $P_K$ is a projection of $M$ onto $K$ in Algorithm 6, we obtain a method for approximating a point in $VIP(G, K)$. Under this settings we have the following strong convergence theorem for approximating a solution of the VIP (6).

**Theorem**

Let $f : K \to K$ be a $\kappa$-contraction and $G : K \to M$ be a pseudomonotone operator satisfying conditions V1-V3, then the sequence $\{x_n\}$ generated by Algorithm 6 converges strongly to an element $p = P_{VIP(G, K)}f(p)$. 

S. Reich and O.K. Oyewole

Equilibrium problem
Application to Convex minimization problem

\[
\begin{aligned}
\min_{x \in K} \phi(x)
\end{aligned}
\]  

(7)

where \( \phi \) is a proper lower semicontinuous convex function of \( M \) onto \((−∞, +∞]\) such that \( M \) is contained in the domain of \( \phi \). That is \( K \subset \text{dom} \phi = \{x \in M : \phi(x) < +\infty\} \). The set of solution of COP (7) is denoted by \( \text{COP}(\phi, K) \). Suppose we define the bifunction \( g : K \times K \to \mathbb{R} \) by \( g(x, y) = \phi(y) - \phi(x) \), then \( g(x, y) \) satisfies the condition (A1)-(A4) and \( \text{COP}(\phi, K) = \text{Sol}(g, K) \). Let \( \text{Prox}_{\lambda \phi} \) be the proximal operator of the function \( \phi \) of parameter \( \lambda > 0 \) and \( \nabla \phi \) is the gradient of \( \phi \). Using the term \( \text{Prox}_{\lambda \phi}(\exp_x(-\lambda \nabla \phi(x))) \) in place of \( \arg \min_{y \in M} \left\{ g(x, y) + \frac{1}{2\lambda_n} d^2(x, y) \right\} \) in Algorithm 6, we obtain a method for minimizing the function \( \phi \).
Numerical example

We consider an extension of the Nash equilibrium model introduced in [16, 21]. In this problem, the bifunction $g : K \times K \rightarrow \mathbb{R}$ is given by

$$g(x, y) = \langle Px + Qy + p, y - x \rangle.$$

Let $M$ be given by Space 2 and $K \subset M$ be given by

$$K = \{ x = (x_1, x_2, \cdots, x_m) : 1 \leq x_i \leq 100, \; i = 1, 2, \cdots, m \},$$

$x, y \in K$, $p = (p_1, p_2, \cdots, p_m)^T \in \mathbb{R}^m$ is chosen randomly with elements in $[1, m]$. The matrices $P$ and $Q$ are two square matrices of order $m$ such that $Q$ is symmetric positive semidefinite and $Q - P$ is negative semidefinite. It is known (see [21]) that $g$ is pseudomonotone, satisfies (A2) with Lipschitz constant $c_1 = c_2 = \frac{1}{2} \| Q - P \|$ (see [39, Lemma 6.2]).
Table: Computation result for Example ??.

<table>
<thead>
<tr>
<th>m</th>
<th>No of Iter.</th>
<th>CPU time (sec)</th>
<th>Algorithm 6</th>
<th>[17, Algorithm 1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>23</td>
<td>0.0013</td>
<td>39</td>
<td>2.9229</td>
</tr>
<tr>
<td>30</td>
<td>23</td>
<td>0.0130</td>
<td>43</td>
<td>3.6771</td>
</tr>
<tr>
<td>50</td>
<td>41</td>
<td>0.0050</td>
<td>53</td>
<td>5.8712</td>
</tr>
<tr>
<td>60</td>
<td>35</td>
<td>0.0050</td>
<td>40</td>
<td>5.8712</td>
</tr>
</tbody>
</table>
The following are some remarks about the current result compared to the existing methods and results in the literature.

(i) The method in this paper uses an adaptive step-size which is allowed to increase from iteration to iteration as against the method in [17] which is monotonically decreasing and the method of [1, 30, 31] which relies on the Lipschitz condition of the bifunction. It is known that the Lipschitz constants can be difficult to estimate which thus affects the efficiency of the method.

(ii) The use of the inertial technique makes the convergence of our faster than the method used in [1, 17].
(iii) With the viscosity method, we obtained a strong convergence theorem which makes our result desirable over the results of Ali-Akbari and Fan et al. [1, 18]. We note that the control parameter of the viscosity step of the method is only required to be non-summable unlike the parameters in [15, 2] which requires an extra condition that the difference between successive parameters be summable. Finally,

(iv) Our result is obtained in the framework of Hadamard manifold unlike the results of [28], [31] and [39] which were obtained in the real Hilbert spaces.
Thanks for your time.


C. Li, and J.C. Yao, Variational inequalities for set-valued vector fields on Riemannian manifolds: convexity of the


T.T.V. Nguyen, J.J. Strodiot and V.H. Nguyen, Hybrid methods for solving simultaneously an equilibrium problem and


