Proximal Methods with Bregman Distances to Solve VIP on Hadamard manifolds*

E. A. Papa Quiroz†**and P. Roberto Oliveira**

Universidad Nacional del Callao
Universidad Nacional Mayor de San Marcos
Lima, Peru
erikpapa@gmail.com

PESC-COPPE
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
poliveir@cos.ufrj.br

Abstract

We present an extension of the proximal point method with Bregman distances to solve Variational Inequality Problems (VIP) on Hadamard manifolds (simply connected finite dimensional Riemannian manifold with nonpositive sectional curvature).

Under some natural assumption, as for example, the existence of solutions of the (VIP) and the monotonicity of the multivalued vector field, we prove that the sequence of the iterates given by the method converges to a solution of the problem.

Keywords: Proximal point methods, Hadamard manifolds, Bregman distances, Variational inequality problems, monotone vector field.

*This paper was developed during a Pos Doctoral visit of the first author to PESC-COPPE Federal University of Rio de Janeiro during the months of January to March 2010, May-June and December 2011

†Corresponding author: E.A. Papa Quiroz, telephone: 51(1)5685537//Mobil Phone: 51(1)992332681

1 Introduction

The Variational Inequality Problem (VIP), from now on, is a mathematical model that embraces diverse applications that arise naturally in science and engineering. Particular cases of these problems are: optimization problems, boundary value problems of Partial Differentiable Equations (PDEs), economic equilibrium problems, traffic network equilibrium problems, among others. We refer to reader to the two volumes monography of Facchinei and Pang [6] for the current state of the art on the theory and algorithms for (VIP) and their existence bibliography.

In this paper we consider the (VIP) define now on Hadamard manifolds (complete simply connected finite-dimensional Riemannian manifolds of non positive sectional curvature): given a nonempty closed convex set X in a Hadamard manifold M, find $x^* \in X$ and $v^* \in V(x^*)$ such that:

$$(VIP) \quad \langle v^*, \exp_{x^*}^{-1} x \rangle \geq 0, \quad \forall x \in X,$$

where $V : M \rightrightarrows TM$ is a multivalued vector field on M, TM is the tangent bundle of the manifold M and $\exp_{x^*}^{-1}$ is the inverse of the exponential map $\exp_{x^*}$. Obviously, when M is the Euclidean space this problem is the classical finite dimensional (VIP).

A motivation to study this subject is the work of Németh [10] who, motivated by the idea of solve boundary value problems on manifolds, proposed the study of (VIP) on Riemannian manifolds. Another one is that using (VIP) on Hadamard manifolds we can solve in particular non convex minimization problems and non monotone singularity problems in Euclidean spaces respectively, that is, if we can transform those problems into convex minimization and monotone singularity problems on Hadamard manifolds respectively, we can use the algorithms introduced on those manifolds.

Another motivation, is that the class of Hadamard manifolds is the natural motivation to study more general spaces of nonpositive curvature such as, for example, Hadamard (also called CAT(0)) and Alexandrov spaces. Observe that spaces of nonpositive curvature play a significant role in many areas: Lie group theory, combinatorial and geometric group theory, dynamical system, harmonic maps and vanishing theorems, geometric topology, Kleinian group theory and Theichmüller theory, see the books [1, 2, 5, 9] for details.

The first study of the (VIP) on Hadamard manifolds is due to Németh [10]. In that paper, the author generalized to Hadamard manifolds some basic existence and uniqueness theorems of the classical theory of variational inequalities on Euclidean spaces. Those results recently have been extended to finite dimensional Riemannian manifolds by Shu-Long Li et al. [8]. The particular problem of finding zeros of monotone vector fields using the proximal point method with riemannian distance has been studied by Da Cruz Neto et al. [3] and Chong Li, et al. [7].

The proximal point method to solve the minimization problem $\min\{f(x) : x \in M\}$ generates a sequence $\{x^k\}$ given by $x^0 \in M$, and

$$x^k \in \mathtt{arg}\min\{f(x) + (\lambda_k/2)d^2(x, x^{k-1}) : x \in M\}, \tag{1.1}$$

where λ_k is a certain positive parameter and d is the Riemannian distance in M. It is well known, see Ferreira and Oliveira, [11], that if M is a Hadamard manifold, f is convex in (1.1) and $\{\lambda_k\}$ satisfies $\sum_{k=1}^{+\infty}(1/\lambda_k) = +\infty$, then $\lim_{k\to\infty} f(x^k) = \inf\{f(x) : x \in M\}$. Furthermore, if the optimal set is nonempty, we obtain that $\{x^k\}$ converges to an optimal solution of the problem.

Recently, Papa Quiroz and Oliveira [13] has been generalized the proximal point method for solving quasiconvex minimization problems by using the following iteration

$$0 \in \widehat{\partial}\left(f(.) + (\lambda_k/2)d^2(., x^{k-1})\right)(x^k) \tag{1.2}$$

where $\widehat{\partial}$ is the regular subdifferential on Hadamard manifolds. Of course both (1.1) and (1.2) are equivalent when f is convex in M, but in the quasiconvex case these iterative schemes are quite different in nature.

For the problem of finding a singularity of a multivalued vector field on a Hadamard manifold, find $\bar{x} \in M$ such that $0 \in V(x)$, Chong Li et al. [7] extended the proximal point method using the iteration:

$$0 \in V(x^{k+1}) - \lambda_k \exp^{-1}_{x_{k+1}} x_k \tag{1.3}$$

The authors proved that if V is a monotone and upper Kuratowski semicontinuous vector field, $V^{-1}(0) \neq \emptyset$ and λ_k satisfies

$$\sup\{\lambda_k : k \geq 0\} < +\infty,$$

and supposing that the sequence $\{x^k\}$ generated by (1.3) is well defined, then $\{x^k\}$ converges to a singularity of V. Furthermore, if $dom(V) = M$ and V is maximal monotone, then the sequence is well defined. Also these authors given some applications in minimization problems with constraints, minimax problems and (VIP) with univalued vector field.

In the present work we deal with the generalization of the well known proximal point method using Bregman distances from Euclidean spaces to Hadamard manifolds to solve (VIP) with multivalued vector field. Our algorithm, given $x^{k-1} \in int(X) \cap S$, find a point $x^k \in X \cap \bar{S}$ by the iteration:

$$0 \in V_k(x^k) := \left(V(\cdot) + \lambda_k \partial D_h(\cdot, x^{k-1})\right)(x^k) \tag{1.4}$$

where X is a closed convex set of M, S is a convex open set of M and D_h is the Bregman distance with zone S, see Definition 4.2 of Section 4.

Observe that our approach has some advantages on the classical proximal point method (1.3). One of them is that we can consider arbitrary Bregman functions, another one is that to find $\{x^k\}$ in each iteration is not need to use the normal cone vector field as it was considered in [7] and therefore, in our opinion, the introduced algorithm is more practical to solve general (VIP).

The paper is organized as follows. In section 2 we give some results on Riemannian geometry that we will use along the paper. In section 3, we present the Variational Inequality Problems on Hadamard manifolds and we give some examples of problems which can be expressed as (VIP) on Hadamard manifolds. In Section 4 the definitions of Bregman function and Bregman distances are introduced and some necessary properties are given. Section 5, we introduce the proximal point algorithm with Bregman distances to solve the (VIP). In Section 6 we analise the global convergence of the algorithm. Finally in Section 7 we give our conclusions.

2 Some Results on Riemannian Geometry

In this paper we give some basic properties and notation of Riemannian manifolds that we are going to use, we refer the reader to do Carmo [4] and Sakai [**?**] for details.

Let M be a differential manifold with finite dimension n. We denote by T_xM the tangent space of M at x and $TM = \bigcup_{x \in M} T_xM$. T_xM is a linear space and has the same dimension of M. Because we restrict ourselves to real manifolds, T_xM is isomorphic to $I\!R^n$. If M is endowed with a Riemannian metric g, then M is a Riemannian manifold and we denote it by (M, G) or only by M when no confusion can arise, where G denotes the matrix representation of the metric g. The inner product of two vectors $u, v \in T_xM$ is written as $\langle u, v\rangle_x := g_x(u, v)$, where g_x is the metrics at point x. The norm of a vector $v \in T_xM$ is set by $||v||_x := \langle v, v\rangle_x^{1/2}$. If there is no confusion we denote $\langle , \rangle = \langle , \rangle_x$ and $||.|| = ||.||_x$. The metrics can be used to define the length of a piecewise smooth curve $\alpha : [t_0, t_1] \to M$ joining $\alpha(t_0) = p'$ to $\alpha(t_1) = p$ through $L(\alpha) = \int_{t_0}^{t_1} \|\alpha'(t)\|_{\alpha(t)} dt$. Minimizing this length functional over the set of all curves we obtain a Riemannian distance $d(p', p)$ which induces the original topology on M.

Given two vector fields V and W in M, the covariant derivative of W in the direction V is denoted by $\nabla_V W$. In this paper ∇ is the Levi-Civita connection associated to (M, G). This connection defines an unique covariant derivative D/dt, where, for each vector field V, along a smooth curve $\alpha : [t_0, t_1] \to M$, another vector field is obtained, denoted by DV/dt. The parallel transport along α from $\alpha(t_0)$ to $\alpha(t_1)$, denoted by P_{α,t_0,t_1}, is an application $P_{\alpha,t_0,t_1} : T_{\alpha(t_0)}M \to T_{\alpha(t_1)}M$ defined by $P_{\alpha,t_0,t_1}(v) = V(t_1)$ where V is the unique vector field along α so that $DV/dt = 0$ and $V(t_0) = v$. Since ∇ is a Riemannian connection, P_{α,t_0,t_1} is a linear isometry, furthermore $P^{-1}_{\alpha,t_0,t_1} = P_{\alpha,t_1,t_0}$ and $P_{\alpha,t_0,t_1} = P_{\alpha,t,t_1} P_{\alpha,t_0,t}$, for all $t \in [t_0, t_1]$. A curve $\gamma : I \to M$ is called a geodesic if $D\gamma'/dt = 0$.

A Riemannian manifold is complete if its geodesics are defined for any value of $t \in \mathbb{R}$. Let $x \in M$, the exponential map $\exp_x : T_xM \to M$ is defined as $\exp_x(v) = \gamma(1)$, where γ is the geodesic such that $\gamma(0) = x$ and $\gamma'(0) = v$. If M is complete, then $\exp_x$ is defined for all $v \in T_xM$. Besides, there is a minimal geodesic (its length is equal to the distance between the extremes).

Given the vector fields X, Y, Z on M, we denote by R the curvature tensor defined by $R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z$, where $[X, Y] := XY - YX$ is the Lie bracket. Now, the sectional curvature as regards X and Y is defined by

$$K(X, Y) = \frac{\langle R(X, Y)Y, X\rangle}{\|X\|^2\|Y\|^2 - \langle X, Y\rangle^2}.$$

The gradient of a differentiable function $f : M \to \mathbb{R}$, $\mathtt{grad}f$, is a vector field on M defined through $df(X) = \langle \mathtt{grad}f, X\rangle = X(f)$, where X is also a vector field on M.

The complete simply-connected Riemannian manifolds with nonpositive curvature are called *Hadamard manifolds.*

Theorem 2.1 *Let M be a Hadamard manifold. Then M is diffeomorphic to the Euclidian space $\mathbb{R}^n$, $n = \dim M$. More precisely, at any point $x \in M$, the exponential mapping $\exp_x : T_xM \to M$ is a global diffeomorphism.*

Proof. See Sakai, [15], Theorem 4.1, page 221. ∎

A consequence of the preceding theorem is that Hadamard manifolds have the property of uniqueness of geodesic between any two points. Another useful property is the following: let $[x, y, z]$ be a geodesic triangle, which consists of *vertices* and the geodesics joining them. We have:

Theorem 2.2 *Given a geodesic triangle $[x, y, z]$ in a Hadamard manifold, it holds that:*

$$d^2(x, z) + d^2(z, y) - 2\langle \exp_z^{-1} x, \exp_z^{-1} y\rangle \leq d^2(x, y),$$

where $\exp_z^{-1}$ denotes the inverse of $\exp_z$.

Proof. See Sakai, [15], Proposition 4.5, page 223. ■

Definition 2.1 *Let M be a Hadamard manifold. A subset A is said to be convex in M if given $x, y \in A$, the geodesic curve $\gamma : [0,1] \to M$ such that $\gamma(0) = x$ and $\gamma(1) = y$ verifies $\gamma(t) \in A$, for all $t \in [0,1]$.*

Definition 2.2 *Let M be a Hadamard manifold and $f : M \to I\!R \cup \{+\infty\}$ be a proper function. f is called convex if for all $x, y \in M$ and $t \in [0,1]$, it holds that*

$$f(\gamma(t)) \leq tf(y) + (1-t)f(x),$$

for the geodesic curve $\gamma : [0,1] \to M$ such that $\gamma(0) = x$ and $\gamma(1) = y$. When the preceding inequality is strict, for $x \neq y$ and $t \in (0,1)$, the function f is called strictly convex.

Theorem 2.3 *Let $f : M \to I\!R \cup \{+\infty\}$ be a proper function with $\mathtt{dom} f$ convex in a Hadamard manifold M. Function $f : M \to I\!R$ is convex in M if and only if $\forall\, x, y \in M$ and $\gamma : [0,1] \to M$ (the geodesic joining x to y) the function $f(\gamma(t))$ is convex in $[0,1]$.*

Proof. See [16], page 61, Theorem 2.2. ■

A function $f : M \to I\!R \cup \{+\infty\}$ is called concave if $-f$ is convex. Furthermore, if f is both convex and concave then f is said to be linear affine on M. It can be proved that, a twice differentiable function f on an open convex set A is linear affine if and only if $\langle H_x^f(v), v\rangle_x = 0$, for all $x \in A$ and $v \in T_xM$. Of fact, $\langle H_x^f(v), v\rangle_x = 0$, if and only if $\langle H_x^f(v), v\rangle_x \geq 0$ and $\langle H_x^f(v), v\rangle_x \leq 0$, if and only if f is convex and concave. In other words, f is linear affine if and only if the vector field $\mathtt{grad} f$ is parallel.

Proposition 2.1 *Let M be a Hadamard manifold and $h : M \to I\!R$ a differentiable function. Let $y \in M$, $v \in T_yM$ and define $g : M \to I\!R$ such that*

$$g(x) = \langle v, \exp_y^{-1} x\rangle_y,$$

for $x \in M$. Then the following statements are true:

i. *$\mathtt{grad} g(x) = P_{\gamma,0,1}v$, where $\gamma : [0,1] \to M$ is the geodesic curve such that $\gamma(0) = y$ and $\gamma(1) = x$.*

ii. *g is an affine function in M.*

Proof. Simmilar to Papa Quiroz and Oliveira, [12], Proposition 3.4. ■

Definition 2.3 *Let $f : M \to I\!R \cup \{+\infty\}$ be a proper convex function. Given $x \in \mathtt{dom} f$, we say that $s \in T_xM$ is a subgradient of f at x if*

$$f(y) \geq f(x) + \langle s, \exp_x^{-1} y\rangle, \forall y \in M$$

The set of all the subgradients of f at x, denoted by $\partial f(x)$, is called the Fecnchel subdifferential of f at x, that is,

$$\partial f(x) := \{s \in T_xM : f(y) \geq f(x) + \langle s, \exp_x^{-1} y\rangle, \forall y \in M\}.$$

Lemma 2.1 *If $p : M \to I\!R \cup \{+\infty\}$ is differentiable at $\bar{x}$ then $\partial p(\bar{x}) = \{\mathtt{grad} p(\bar{x})\}$*

Proof. Since p is differentiable at $\bar{x}$ we have

$$p(\gamma(t)) = p(\bar{x}) + t\langle \mathtt{grad}p(\bar{x}), \gamma'(0)\rangle + o(t), \tag{2.5}$$

for all geodesic γ such that $\gamma(0) = \bar{x}$, where $\lim_{t\to 0} o(t)/t = 0$. Let $s \in \partial p(\bar{x})$, then

$$p(\gamma(t)) \geq p(\bar{x}) + t\langle s, \gamma'(0)\rangle, \tag{2.6}$$

for all geodesic γ such that $\gamma(0) = \bar{x}$, where $\lim_{t\to 0} \theta(t)/t = 0$. From (2.5) and (2.6) we obtain

$$0 \geq \langle s - \mathtt{grad}p(\bar{x}), \gamma'(0)\rangle - \frac{o(t)}{t}$$

for all geodesic γ such that $\gamma(0) = \bar{x}$. Taking in particular γ such that $\gamma(0) = \bar{x}$ and $\gamma'(0) = s - \mathtt{grad}f(\bar{x})$, and taking $t \to 0$, we have

$$s = \mathtt{grad}p(\bar{x}). \quad \blacksquare$$

Definition 2.4 *Let M be a Hadamard manifold. A multivalued set vector field V from M to TM, denoted by $V : M \rightrightarrows TM$ is an application from M to TM such that $V(x) \subset T_xM$. The set*

$$dom(V) = \{x \in M : V(x) \neq \emptyset\}$$

is called the domain of V.

Definition 2.5 *We say that the multivalued vector field $V : M \rightrightarrows TM$ is monotone if for all $x, y \in M$ the following condition is satisfied*

$$\langle P_{\gamma,0,1}v - w, \exp_x^{-1} y\rangle \geq 0, \tag{2.7}$$

for all $v \in V(y)$ and all $w \in V(x)$, where γ is the (unique) geodesic such that $\gamma(0) = y$ and $\gamma(1) = x$.

Next, we introduce the definition of maximal vector fields and give a fundamental result which will be very important to the convergence of the proposed method when $\overline{dom(V)} \subseteq int(X) \cap S$.

Definition 2.6 *A multivalued vector field $V : M \rightrightarrows TM$ is maximal monotone if*

i) *V is monotone*

ii) *For all V' monotone such that $V(x) \subset V'(x)$ for all $x \in M$, it holds $V = V'$.*

The following lemma shows that maximal monotone vector fields on Hadamard manifolds are closed.

Lemma 2.2 *If $\lim_{k\to+\infty} z^k = \bar{z}$, $V : M \to TM$ is maximal monotone and $\lim_{k\to+\infty} P_{\gamma_k,0,1}y^k = \bar{y}$, where γ_k is the geodesic such that $\gamma_k(0) = z^k$ and $\gamma_k(1) = \bar{z}$, where $y^k \in V(z^k)$, then $\bar{y} \in V(\bar{z})$.*

Proof. Define

$$V'(z) = \begin{cases} V(z), & z \neq \bar{z} \\ V(\bar{z}) \cup \{\bar{y}\}, & z = \bar{z} \end{cases}$$

We claim that V' is monotone. We need to show

$$\langle P_{\gamma,0,1} y - y', \exp^{-1}_{z'} z\rangle \geq 0, \forall z, z \in M; \forall y \in V'(z); \forall y' \in V'(z'),$$

where γ is the geodesic such that $\gamma(0) = z$ and $\gamma(1) = z'$. By the monotonocity of V it suffices to check the above inequality only for $y' = \bar{y}$, $z' = \bar{z}$. From monotonicity of V we have

$$\langle P_{\alpha_k,0,1} y - y^k, \exp^{-1}_{z^k} z\rangle \geq 0, \forall z \in M; \forall y \in V(z), \tag{2.8}$$

where α_k is the geodesic such that $\alpha_k(0) = z$ and $\alpha_k(1) = z^k$. Taking limits in (2.8) we obtain

$$\langle P_{\gamma,0,1} y - \bar{y}, \exp^{-1}_{\bar{z}} z\rangle \geq 0, \forall z \in M; \forall y \in V(z).$$

So V' is monotone. Since $V(x) \subset V'(x)$, for all $x \in M$ and V is monotome we conclude that $V = V'$. In particular we have $V(\bar{z}) = V'(\bar{z}) = V(\bar{z}) \cup \{\bar{y}\}$, that is, $\bar{y} \in V(\bar{z})$ and the lemma is proved. ■

3 VIP on Hadamard Manifolds

In this paper we are interested in solving the (VIP) on Hadamard manifolds: given a nonempty closed convex set X on a Hadamard manifold M, find $x^* \in X$ and $v^* \in V(x^*)$ such that

$$\langle v^*, \exp^{-1}_{x^*} x\rangle \geq 0, \quad \forall x \in X, \tag{3.9}$$

where $V : M \rightrightarrows TM$ is a multivalued vector field on M.
In an attempt to motivate our work we will give some examples of problems which can be expressed as (VIP) on Hadamard manifolds.

3.1 Convex minimization problems on Hadamard manifolds.

Let $f : M \to I\!R$ be a convex function and X be a nonempty closed convex set on a Hadamard manifold M. Consider the following optmization problem:

$$\min\{f(x) : x \in X\}.$$

It can be easily proved, using Proposition 5.4 of [7], that this problem is equivalent to (3.9) for $V = \partial f$.

3.2 Singularity Problems on Hadamard manifolds

Let M be a Hadamard manifold and consider $V : M \rightrightarrows TM$ be a multivalued monotone vector field. The singularity problem is to find a point $\bar{x} \in M$ such that

$$0 \in V(\bar{x}).$$

It can be show that this problem is equivalent to (3.9) when $X = M$.

3.3 Nash Equilibrium Problems on Hadamard Manifolds:

In a general noncooperative game, there are n players each of whom has a certain cost function and strategy set that may depend on the other player's actions. We assume that player i's strategy set is X_i, which is a subset of the Hadamard manifold M^{n_i} and is independt of the other player's strategies. Player i's cost function $\mu_i(x)$ depends of all players strategies, which are described by the point $x = (x^1, x^2, ..., x^n)$, where $x^i \in M^{n_i}$ for $i = 1, ..., n$. Player i's problem is to determine, for each fixed but arbitrary tuple $\widetilde{x}^i \equiv (x^j : j \neq i)$ of the other player's strategies an optimal strategy x^i that solves the cost minimization problem in the variable y^i :

$$\begin{cases} min\ \mu_i(y_i, \widetilde{x}^i) \\ s.a: \\ \qquad y_i \in X_i \end{cases}$$

we denote the solution set of this optimization problem by $S_i(\widetilde{x}^i)$; note the dependence of this set on the tuple $\widetilde{x}^i$. A slight abuse of notation ocurrs in the objective function $\mu_i(y_i, \widetilde{x}^i)$; it is understood that this mean the function μ_i evaluated at the points whose $j-th$ subvector is x^j for $j \neq i$ and whose $i-$th subvector is y^i.

A Nash equilibrium on a Hadamard manifold is a tupla of strategies $x = (x^i : i = 1, 2, ..., n)$ with the property that for each i, $x^i \in S(\widetilde{x}^i)$.

The following result gives a set of sufficient conditions under which a Nash equilibrium can be obtained by solving a (VIP).

Proposition 3.1 *Let each X_i be a closed convex subset of a Hadamard manifold $M_i^{n_i}$. Suppose that for each fixed tuple $\widetilde{x}^i$, the function $\mu_i(., \widetilde{x}^i) : M \to I\!R$ is convex and continuously differentiable. Then a tuple $x = (x^i : i = 1, ..., n)$ is a Nash equilibrium if, and only if, x solves the (VIP) where $X = \prod_{i=1}^n X_i$ and $V(x) = ((grad_{x^i}\, \mu_i(x))_{i=1}^n)$.*

Proof. By convexity and the minimum principle, if x is a Nash Equilibrium then for each $i = 1, ..., n$

$$\left\langle grad_{x^i}\, (\mu_i(x)), \exp_{x^i}^{-1} y^i \right\rangle \geq 0,\ \forall y^i \in X_i.$$

Thus, we obtain,

$$\langle V(x), \exp_x^{-1} y \rangle = \sum_{i=1}^{n} \left\langle grad_{x^i}\, (u_i(x)), \exp_{x^i}^{-1} y^i \right\rangle \geq 0, \forall y \in X.$$

Conversely, if x solves the (VIP) where $X = \prod_{i=1}^n X_i$ and $V(x) = ((grad_{x^i}\, \mu_i(x))_{i=1}^n)$, then

$$\langle V(x), \exp_x^{-1} y \rangle \geq 0, \forall y \in X.$$

In particular, for each $i = 1, ..., n$, let y be the tuple whose $j - th$ subvector is equal to x^j, for $j \neq i$, and $i-$th subvector is equal to y^i, is an arbitrary element of the set X^i. The above inequality then becomes

$$\left\langle grad_{x^i}\, (u_i(x)), \exp_{x^i}^{-1} y^i \right\rangle \geq 0,\ \forall y^i \in X_i.$$

■

4 Bregman Distances on Hadamard Manifolds

Let M be a Hadamard manifold and S a nonempty open convex set of M with a topological closure $\bar{S}$. Let $h : \bar{S} \to I\!R$ be a differentiable function in S. Define the funcion $D_h(.,.) : \bar{S} \times S \to I\!R$ so that

$$D_h(x, y) := h(x) - h(y) - \langle \mathtt{grad} h(y), \exp_y^{-1} x\rangle_y. \tag{4.10}$$

Let us adopt the following notation for the partial level sets of D_h. For $\alpha \in I\!R$, take

$$\Gamma_1(\alpha, y) := \{x \in \bar{S} : D_h(x, y) \leq \alpha\},$$

$$\Gamma_2(x, \alpha) := \{y \in S : D_h(x, y) \leq \alpha\}.$$

Definition 4.1 *Let M be a Hadamard manifold. A real function $h : M \to I\!R$ is called a* Generalized Bregman function*, denoted by $h \in \mathcal{GB}$, if there exists a nonempty open convex set S so that*

a. *h is continuous on $\bar{S}$;*

b. *h is strictly convex on $\bar{S}$;*

c. *h is continuously differentiable in S;*

d. *For all $\alpha \in I\!R$ the partial level sets $\Gamma_1(\alpha, y)$ and $\Gamma_2(x, \alpha)$ are bounded for every $y \in S$ and $x \in \bar{S}$, respectively.*

In this case D_h is called generalized Bregman distance *from x to y.*

Definition 4.2 *Let M be a Hadamard manifold. A real function $h : M \to I\!R$ is called a* Bregman function*, denoted by $h \in \mathcal{B}$, if there exists a nonempty open convex set S such that h satisfies the conditions of the aforementioned definition and, also satisfies:*

e. *If $\lim_{k\to+\infty} y^k = y^* \in \bar{S}$, then $\lim_{k\to+\infty} D_h(y^*, y^k) = 0$, and*

f. *If $\lim_{k\to+\infty} D_h(z^k, y^k) = 0$, $\lim_{k\to+\infty} y^k = y^* \in \bar{S}$ and $\{z^k\}$ is bounded, then $\lim_{k\to+\infty} z^k = y^*$.*

In this case D_h is called Bregman distance *from x to y.*

From the above definitions we obtain, obviously, that $\mathcal{B} \subseteq \mathcal{GB}$ and the equality is satisfied when $S = M$. In both definitions, the set S is called the *zone* of the function h. Some examples of Bregman distance for different Riemannian manifolds have been provided by [12], Section 8.

Lemma 4.1 *Let $h \in \mathcal{GB}$ with zone S. Then*

i. *$\mathbf{grad}\, D_h(., y)(x) = \mathbf{grad}\, h(x) - P_{\gamma,0,1}\mathbf{grad}\, h(y)$, for all $x, y \in S$, where $\gamma : [0, 1] \to M$ is the geodesic curve such that $\gamma(0) = y$ and $\gamma(1) = x$.*

ii. *$D_h(., y)$ is strictly convex on $\bar{S}$ for all $y \in S$.*

iii. *For all $x \in \bar{S}$ and $y \in S$, $D_h(x, y) \geq 0$ and $D_h(x, y) = 0$ if and only if $x = y$.*

Proof. Analogous to Lemma 4.1 of [12]. ■

Note that D_h is not a distance in the usual sense of the term. In general, the triangular inequality is not valid, as the symmetry property.
From now on, we use the notation $\mathtt{grad}\, D_h(x, y)$ to mean $\mathtt{grad}\, D_h(., y)(x)$. So, if γ is the geodesic curve so that $\gamma(0) = y$ and $\gamma(1) = x$, from Lemma 4.1, i, we obtain

$$\mathtt{grad} D_h(x, y) = \mathtt{grad}\, h(x) - P_{\gamma,0,1}\mathtt{grad}\, h(y).$$

Lemma 4.2 *Let $h \in \mathcal{GB}$ with zone S and $y \in S$. Suppose that $z \in S$, then, the function*

$$G(x) := D_h(x, y) - D_h(x, Py)$$

is a linear affine on $\bar{S}$.

Proof. From (4.10) $G(x) = h(z) - h(y) + \langle \mathtt{grad} h(z), \exp_z^{-1} x\rangle_z - \langle \mathtt{grad} h(y), \exp_y^{-1} x\rangle_y$. Due to the affine linearity of functions $\langle \mathtt{grad} h(z), \exp_z^{-1} x\rangle_z$ and $\langle \mathtt{grad} h(y), \exp_y^{-1} x\rangle_y$ at x (See Proposition 2.1, **ii**) the result follows. ■

Proposition 4.1 *Let $h \in \mathcal{GB}$ with zone S. For all $y, z \in S$ and $x \in \bar{S}$ we have*

$$\langle \mathit{grad} D_h(z, y), \exp_z^{-1} x\rangle = D_h(x, y) - D_h(x, z) - D_h(z, y).$$

Proof. Let $\gamma : [0, 1] \to M$ be the geodesic curve so that $\gamma(0) = z$ and $\gamma(1) = x$. Due to Lemma 4.2 the function $G(x) = D_h(x, y) - D_h(x, z)$ is linear affine on $\bar{S}$. That is, G is both convex and concave. Then, from Theorem 2.3 we have $G(\gamma(t)) = tG(x) + (1 - t)G(z)$, which gives,

$$\frac{1}{t}(D_h(\gamma(t), y) - D_h(z, y)) - \frac{1}{t}(D_h(\gamma(t), z) - D_h(z, z)) = (D_h(x, y) - D_h(x, z) - D_h(z, y))$$

where we took into account that $D_h(z, z) = 0$. Taking limit when $t \to 0$ we obtain

$$\langle \mathtt{grad} D_h(z, y), \exp_z^{-1} x\rangle = D_h(x, y) - D_h(x, z) - D_h(z, y). \quad ■$$

Definition 4.3 *h is said be a boundary coercive function if for each $\{x^k\} \subset S$ convergent to some point x with $x \in Bound(S)$ (boundary of S) we have*

$$\lim_{k\to\infty} \langle \mathit{grad} h(x^k), \exp_{x^k}^{-1} y\rangle = -\infty$$

Lemma 4.3 *Let $h \in \mathcal{GB}$ with zone S and boundary coercive. Then, for each $y \in S$ we obtain*

$$\partial_x D_h(x, y) = \begin{cases} \mathit{grad}\, h(x) - P_{\gamma,0,1}\mathit{grad}\, h(y), & x \in S \\ \emptyset, & x \notin S \end{cases} \tag{4.11}$$

Where $\partial_x D_h(x, y)$ is the Fenchel subdifferential of $D_h(., y)$ at x.

Proof. If $x \in S$, it holds by Lemma 2.1 and Lemma 4.1. It is also clear that $\partial_x D_h(x, y) = \emptyset$ for $x \notin \bar{S}$. Therefore, we need consider the case when $x \in Bound(S)$. Take $x \in Bound(S)$ and suppose there exists $v \in \partial_x D_h(x, y)$. Let $\{\zeta_k\} \subset (0, 1)$ be a sequence such that $\lim_{k\to\infty} \zeta_k = 0$, and we define

$$x^k := \exp_x\left(\zeta_k \exp_x^{-1} y\right).$$

This implies that

$$\zeta_k \langle v, \exp_x^{-1} y\rangle = \langle v, \exp_x^{-1} x^k\rangle. \tag{4.12}$$

It is easy to check that $x^k \in S$. Besides, As $v \in \partial_x D_h(x, y)$ we have

$$D_h(x^k, y) \geq D_h(x, y) + \langle v, \exp_x^{-1} x^k\rangle$$

implying that

$$\langle v, \exp_x^{-1} x^k\rangle \leq D_h(x^k, y) - D_h(x, y) \tag{4.13}$$

Now, by Lemma 4.1, **ii**, $D_h(x, y) > D_h(x^k, y) + \langle \mathtt{grad} D_h(x^k, y), \exp_{x^k}^{-1} x\rangle$, and thus

$$D_h(x^k, y) - D_h(x, y) \quad < \quad -\langle \mathtt{grad} D_h(x^k, y), \exp_{x^k}^{-1} y\rangle. \tag{4.14}$$

From (4.12) and inequalities (4.13) and (4.14) we obtain

$$\zeta_k \langle v, \exp_x^{-1} y\rangle = \langle v, \exp_x^{-1} x^k\rangle < -\langle \mathtt{grad} D_h(x^k, y), \exp_{x^k}^{-1} x\rangle. \tag{4.15}$$

We consider the function

$$g(z) := \langle w, \exp_{x^k}^{-1} z\rangle, \quad \forall\, w \in T_{x^k} M \tag{4.16}$$

which, by Proposition 2.1, **ii**, is a linear affine function, that is, $g(\gamma(t)) = tg(y) + (1-t)g(x), \forall t \in [0, 1]$ where $\gamma : [0, 1] \to M$ is the geodesic curve such that $\gamma(0) = x$ and $\gamma(1) = y$. Taking $t = \zeta_k$ in the previous equality and as $g(x^k) = 0$ we have $g(x) = \frac{\zeta_k}{(\zeta_k - 1)} g(y)$, i.e,

$$\langle w, \exp_{x^k}^{-1} x\rangle = \frac{\zeta_k}{(\zeta_k - 1)} \langle w, \exp_{x^k}^{-1} y\rangle. \tag{4.17}$$

We define,

$$w^k := -\mathtt{grad} D_h(x^k, y) \tag{4.18}$$

Using (4.17) and (4.18) in (4.15), we have

$$\langle v, \exp_x^{-1} y\rangle \quad \leq \quad \frac{1}{\zeta_k - 1} \langle w^k, \exp_{x^k}^{-1} y\rangle \tag{4.19}$$

Observe that $D_h(u, v) + D_h(v, u) = -\langle \mathtt{grad} D_h(v, u), \exp_v^{-1} u\rangle$. Using this identity in (4.19), we have

$$(1 - \zeta_k)\, \langle v, \exp_x^{-1} y\rangle + D_h(x^k, y) \leq -D_h(y, x^k).$$

Taking $k \to \infty$ and using the boundary coercivity property of h and the definition of D_h, we have $\langle v, \exp_x^{-1} y\rangle + D_h(x, y) \leq -\infty$, which is a contradiction. Therefore $\partial_x D_h(x, y) = \emptyset$ for all $x \in \partial S$. ■

4.1 Fejér Convergence with Bregman Distances

Definition 4.4 *Let M be a Hadamard manifold. A sequence $\{y^k\}$ of M is D_h-*Fejér convergent *to a nonempty set $U \subset M$, if*

$$D_h(u, y^{k+1}) \leq D_h(u, y^k),$$

for every $u \in U$.

Theorem 4.1 *Let M be a Hadamard manifold and $h \in \mathcal{GB}$ with zone S. If $\{y^k\}$ is D_h-Fejér convergent to a nonempty set $U \subset M$, then $\{y^k\}$ is bounded. If, furthermore, $h \in \mathcal{B}$ and a cluster point $\bar{y}$ of $\{y^k\}$ belongs to U, then $\{y^k\}$ converges and $\lim_{k\to+\infty} y^k = \bar{y}$.*

Proof. From the above definition

$$0 \le D_h(u, y^k) \le D_h(u, y^0), \tag{4.20}$$

for all $u \in U$. Thus, $y^k \in \Gamma_2(u, \alpha)$ with $\alpha = D_h(u, y^0)$. We can now apply Definition 4.1, **d**, to see that $\{y^k\}$ is bounded.
Let $\bar{y}$ a cluster point of $\{y^k\}$, with $\bar{y} \in U$, then there exists a subsequence $\{y^{k_j}\}$ so that $\lim_{j\to\infty} y^{k_j} = \bar{y}$. From Definition 4.1, **e**, it is true that $\lim_{j\to+\infty} D_h(\bar{y}, y^{k_j}) = 0$. From (4.20), $\{D_h(\bar{y}, y^k)\}$ is a nonincreasing bounded below sequence with a subsequence converging to 0, hence the overall sequence converges to 0, that is,

$$\lim_{k\to+\infty} D_h(\bar{y}, y^k) = 0. \tag{4.21}$$

To prove that $\{y^k\}$ has an unique limit point, let y' be another limit point of $\{y^k\}$. From (4.21) $\lim_{l\to+\infty} D_h(\bar{y}, y^{k_l}) = 0$ with $\lim_{l\to+\infty} y^{k_l} = y'$. Using Definition 4.1, **f**, we have $y' = \bar{y}$. It follows that $\{y^k\}$ cannot have more than one limit point and therefore, $\lim_{k\to+\infty} y^k = \bar{y}$. ■

5 The Algorithm

We are interested in solving the (VIP): given a nonempty closed convex set X of a Hadamard manifold M and $V : M \rightrightarrows TM$ a multivalued vector field, find $x^* \in X$ and $v^* \in V(x^*)$ such that

$$\langle v^*, \exp_{x^*}^{-1} x \rangle \ge 0, \quad \forall x \in X,$$

PBM Algorithm

Let $h \in \mathcal{GB}$ with zone S, as defined in Section 4, such that $X \cap \bar{S} \neq \emptyset$, and let D_h be the function associate to h and defined by (4.10).

Initialization:
Let $\{\lambda_k\}$ be a sequence of positive parameters and an initial point

$$x^0 \in int(X) \cap S. \tag{5.22}$$

Main Steps:
For $k = 1, 2, 3, ...,$ given $x^{k-1} \in int(X) \cap S$
If $0 \in V(x^{k-1})$, then stop.
Otherwise, find $x^k \in X \cap \bar{S}$, such that

$$0 \in V_k(x^k) := \Big(V(\cdot) + \lambda_k \partial D_h(\cdot, x^{k-1})\Big)(x^k) \tag{5.23}$$

where $\partial D_h(\cdot, x^{k-1})$ is the Fenchel subdifferential of $D_h(\cdot, x^{k-1})$
Take $k = k + 1$.

Along the paper we assume the following:

General Assumption (GA): For each $k \in \mathbb{N}$ there exists $x^k \in int(X) \cap S$ satisfying (5.23).

Remark 5.1 *Since that a broad class of Generalized Bregman functions satisfies Assumption* (GA)*, for example boundary coercive Bregman functions, see Lemma 4.3, that assumption is not very restrictive.*

Remark 5.2 *Under assumption* **(GA)***, we have that there exists* $g^k \in V(x^k)$ *such that*

$$g^k = -\lambda_k \textbf{\textit{grad}} D_h(x^k, x^{k-1}).$$

As we are interested in the asymptotic convergence of the method, we also assume that the algorithm not stop in a finite number of iterations, that is, $0 \notin V(x^{k-1})$ for all k, which immediately implies that $x^k \neq x^{k-1}$, for all k.

6 Convergence for Monotone Vector Fields

In this section we also assume the following assumption:

Assumption 1: $X^* \cap \bar{S} \neq \emptyset$, where X^* is the solution set of the (VIP) and S is the zone of $h \in \mathcal{GB}$.

Theorem 6.1 *Under assumptions* **GA**, **1**, $h \in \mathcal{GB}$ *and the monotonicity of* V*, the following inequality holds*

$$D_h(x^*, x^k) \leq D_h(x^*, x^{k-1}) - D_h(x^k, x^{k-1}),$$

for all $x^* \in X^* \cap \bar{S}$*. Thus,* $\{x^k\}$ *is* D_h*-Fejér convergent to* $X^* \cap \bar{S}$*.*

Proof. By assumption GA, $x^k \in intX \cap S$. As V is monotone, taking $y = x^k$ and $x = x^*$ in (2.7) we have

$$\langle P_{\gamma_k,0,1} v^k, \exp^{-1}_{x^*} x^k \rangle \geq 0, \tag{6.24}$$

for all $v^k \in V(x^k)$, where $\gamma_k : [0,1] \to M$ is the geodesic such that $\gamma_k(0) = x^k$ and $\gamma_k(1) = x^*$. Since that the parallel transport is a linear isometry and from above inequality we have $\langle v^k, \exp^{-1}_{x^k} x^* \rangle \leq 0$. Now, from Remark 5.2, taking in particular $v^k = -\lambda_k \mathtt{grad} D_h(x^k, x^{k-1})$, we have that

$$\langle \mathtt{grad} D_h(x^k, x^{k-1}), \exp^{-1}_{x^k} x^* \rangle \geq 0.$$

Taking $z = x^k, x = x^*$ and $y = x^{k-1}$ in Proposition 4.1 we have

$$D_h(x^*, x^k) \leq D_h(x^*, x^{k-1}) - D_h(x^k, x^{k-1}). \tag{6.25}$$

Thus, we obtain $D_h(x^*, x^k) \leq D_h(x^*, x^{k-1})$, since that $D_h(x^k, x^{k-1}) \geq 0$. Then, by Definition 4.4, $\{x^k\}$ is D_h-Fejer convergent to $X^* \cap \bar{S}$. ∎

Corollary 6.1 *Under assumptions of the previous Theorem, the sequence* $\{x^k\}$*, generated by the algorithm, is bounded.*

Proof. It is a consequence of Theorem 4.1. ∎

Proposition 6.1 *Under assumptions* **GA**, **1**, $h \in \mathcal{GB}$ *and the monotonicity of* V*, the following fact are true:*

a) *For all* $x^* \in X^* \cap \bar{S}$, $\{D_h(x^*, x^k)\}$ *converges;*

b) $\lim_{k\to\infty} D_h(x^k, x^{k-1}) = 0$;

c) *Furthermore, if* $h \in \mathcal{B}$ *and* $\lim_{j\to\infty} x^{k_j-1} = \bar{x}$ *then* $\lim_{j\to\infty} x^{k_j} = \bar{x}$.

Proof.
a) From Theorem 6.1 and Lemma 4.1, **iii**, we have $0 \leq D_h(x^*, x^k) \leq D_h(x^*, x^{k-1})$. Thus, $\{D_h(x^*, x^k)\}$ is a nonincreasing sequence bounded from below and hence convergent.
b) Rearranging terms in the inequality (6.25), we have that

$$0 \leq D_h(x^k, x^{k-1}) \leq D_h(x^*, x^{k-1}) - D_h(x^*, x^k).$$

Using item a), we obtain the result.
c) Let $\{x^{k_j}\}$ be a subsequence of $\{x^k\}$ such that $\lim_{j\to\infty} x^{k_j-1} = \bar{x}$. From item b) $\lim_{j\to\infty} D_h(x^{k_j}, x^{k_j-1}) = 0$; As $\{x^{k_j}\}$ is bounded, then, applying Definition 4.2, **f**, we get $\lim_{j\to\infty} x^{k_j} = \bar{x}$. ■

Theorem 6.2 *Under assumptions* **GA**, **1**, $h \in \mathcal{B}$ *and the monotonicity of* V, *and if one of the following condition is satisfied:*

i. T *is maximal and* $\overline{dom(V)} \subseteq int(X) \cap S$;

ii. *There exists* $x^* \in X^* \cap \bar{S}$ *such that* $\lim_{k\to\infty} D_h(x^*, x^k) = 0$

then, the sequence $\{x^k\}$ *converges to a solution of the* VIP.

Proof. From Corollary 6.1, we have that $\{x^k\}$ is bounded then there exists a point $\bar{x} \in X \cap \bar{S}$ and a subsequence $\{x^{k_j}\}$ such that $\lim_{j\to\infty} x^{k_j-1} = \bar{x}$. From Proposition 6.1, **c**, we have that $\lim_{j\to\infty} x^{k_j} = \bar{x}$.
i. As $\overline{dom(V)} \subseteq int(X) \cap S$ and From Remark 5.2 and Proposition 6.1, **c**, we have that $g^{k_j} \in V(x^{k_j})$ and $g^{k_j} \to 0$. Then using Lemma 2.2 we obtain that $0 \in V(\bar{x})$ and thus $\bar{x} \in X^*$. Then using $h \in \mathcal{B}$, theorems 6.1 and 4.1 we obtain the result.
ii. Let $y^j = x^{k_j} \in int(X) \cap S$ and $z^j = x^*$ and from assumption $\lim_{j\to\infty} D_h(z^j, y^j) = 0$. Then, using Definition 4.2 item f., we obtain that $\bar{x} = x^* \in X^*$. Finally using that $h \in \mathcal{B}$ and Theorem 4.1 we obtain that $\{x^k\}$ converges to x^*. ■

7 Conclusion

In this paper we prove the convergence of a proximal point method using Bregman distances to solve (VIP) with multivalued monotone vector fields. To the our knowledge, this is the first attempt to develop a method with generalize distances on Hadamard manifolds. For a computational implementation of the proposed method it is needed to solve the iteration (5.23) using a local algorithm, which only provides an approximate solution. Therefore, we consider that in a future work it is important to analyze the convergence of the proposed algorithm considering now an inexact iteration.

On the other hand, observe that the assumption **ii** of Theorem 6.2 is very strong to obtain the convergence of the proposed method. To remove this assumption is necessary to make a deeper study of the monotone vector fields on Hadamard manifolds. This is a research that we are doing in the working paper [14].

References

[1] W. Ballmann: Lectures on Spaces of Nonpositive Curvature, Birkhäuser Basel (1995).

[2] M.R. Bridson, A. Haefliger: Metric Spaces of Non-Positive Curvature, Springer-Verlag, Berlin (1999).

[3] J.X. Da Cruz Neto, O.P. Ferreira, L.R Lucambio Pï¿½rez, S.Z. Nï¿½meth: Convex- and monotone-transformable mathematical programming problems and a proximal-like point method, J. Glob. Optim. 35(1) (2006) 53-69.

[4] M.P. do Carmo: Riemannian Geometry, Bikhausen, Boston (1992).

[5] P.B. Eberlein, Geometry of Nonpositively Curved Manifolds, University of Chicago Press, Chicago (1996).

[6] F. Facchinei, J.S. Pang: Finite-Dimensional Variational Inequalities and Complementarity Problems. Vol. I and II. Springer Series in Operations Research, Springer-Verlag, NewYork (2003).

[7] Chong Li, G. López, V. Martï¿½n-Mï¿½rquez: Monotone vector fields and the proximal point algorithm on Hadamard manifolds, J. Lond. Math. Soc.(2) 79(3) (2009) 663-683.

[8] Shu-Long Li, Chong Li, Yeong-Cheng Lio, Jen-Chih Yao: Existence of solution for variational inequalities on riemannian manifolds, Nonlinear Anal. 71 (11) (2009) 5695-5706.

[9] J. Jost: Non Positive Curvature: Geometric and Analytic Aspects, Lectures in Mathematics, Base; Boston, Birkhauser (1997).

[10] S.Z. Németh: Variational inequalities on Hadamard manifolds, Nonlinear Anal. 52 (5) (2003) 1491-1498.

[11] O.P. Ferreira, P.R. Oliveira: Proximal point algorithm on riemannian manifolds, Optimization 51(2) (2002) 257-270.

[12] E.A. Papa Quiroz, E.A., P.R. Oliveira: Proximal point methods for quasiconvex and convex functions with Bregman distances on Hadamard manifolds, J. Convex Anal. 16(1) (2009) 49-69.

[13] E.A. Papa Quiroz, E.A., P.R. Oliveira: Full Convergence of the proximal point method for quasiconvex functions on Hadamard manifolds, to appear in ESAIM-COCV-Control, Optimization and Calculus of Variations, 2011.

[14] E.A. Papa Quiroz, E.A., P.R. Oliveira: Maximal monotone vector fields and the PPM with Bregman distances on Hadamard manifolds, working paper.

[15] T. Sakai: Riemannian Geometry, American Mathematical Society, Providence, RI, (1996).

[16] C. Udriste: Convex Function and Optimization Methods on Riemannian Manifolds, Kluwer Academic Publishers, Dordrecht (1994).