Journal of Global Optimization (2006) 35: 53–69
DOI 10.1007/s10898-005-6741-9

Convex- and Monotone-Transformable Mathematical Programming Problems and a Proximal-Like Point Method

J. X. DA CRUZ NETO[1], O. P. FERREIRA[2], L. R. LUCAMBIO PÉREZ[3] and S. Z. NÉMETH[4]
[1]*DM, Universidade Federal do Piauí, Teresina, PI 64049-500, BR (E-mail: jxavier@ufpi.br)*
[2]*IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, BR (E-mail: orizon@mat.ufg.br)*
[3]*IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, BR (E-mail: lrlp@mat.ufg.br)*
[4]*Computer and Automation Institute, Hungarian Academy of Sciences, 1518 Budapest, P.O. Box 63, Hungary (E-mail: snemeth@sztaki.hu)*

(Received 1 October 2003; accepted in revised form 22 April 2005)

Abstract. The problem of finding the singularities of monotone vectors fields on Hadamard manifolds will be considered and solved by extending the well-known proximal point algorithm. For monotone vector fields the algorithm will generate a well defined sequence, and for monotone vector fields with singularities it will converge to a singularity. It will also be shown how tools of convex analysis on Riemannian manifolds can solve non-convex constrained problems in Euclidean spaces. To illustrate this remarkable fact examples will be given.

1. Introduction

Convexity is a sufficient but not necessary condition for many important results in mathematical programming, since there are diverse extensions of the notion of convexity bearing the same properties, e.g., the critical points of pseudo-convex and strictly quasi-convex differentiable functions are global minimizers. Moreover, it is possible to modify numerical methods to solve non-convex optimization problems, e.g., the steepest descent method with a proximal regularization [6] or with Armijo's stepsize [2] generates a sequence that, starting at any point of $\mathbb{R}^n$, converges to a minimizer of a pseudo-convex differentiable function.

Classically, a function is convex if and only if its restriction to each line segment in its domain is convex. This property inspired Ortega and Rheinboldt [11], Avriel [1] and others to introduce the concept of arcwise convex functions. The idea of arcwise convexity is strongly dependent on the way of producing a suitable family of curves. By using the tools of

Riemannian geometry, Rapcsák [15] introduces a modern novel method to investigate arcwise convexity.

Inspired by Rapcsák [15] and Udrişte's [17] geometrical viewpoint, besides some nonconvex problems, we shall consider nonmonotone problems, too. We shall solve them by extending the proximal point algorithm. Let us start with a nonconvex and a nonmonotone problem of the form:

$$\min_{p\in M} f(p), \tag{1}$$

and

$$\text{find } x\in M \text{ such that } X(x)=0, \tag{2}$$

where M is a subset of the Euclidean space $\mathbb{R}^n$, $f\colon M\to\mathbb{R}$ is a function and $X\colon M\to\mathbb{R}^n$ a vector field. By choosing an appropriate Riemannian metric on M, sometimes we can transform problems (1) and (2) into a convex and monotone unconstrained problem on M, respectively, that can be studied by using the intrinsic geometry of M. Since there is an analogy of ideas, we shall use this parallel approach of optimization and singularity problems throughout the paper. In the meantime, note that for a gradient vector field (i.e., a vector field that is the gradient of a function with respect to the metric of M), a singularity problem is equivalent to an optimization problem, and if the gradient vector field is monotone (with respect to the metric of M [8]), it is equivalent to a convex optimization problem (with respect to the metric of M [14]). Bearing this in mind, problem (2) can be viewed as a non-gradient extension of problem (1) considered by Rapcsák [15]. The examples given for problem (1) follow the ideas of Rapcsák and will be presented here for the sake of parallelism between gradient (i.e., optimization problems) and non-gradient singularity problems. Several methods have been extended to Riemannian context to solve problems (1) and (2), see [14, 15, 17–20]. However, solving optimization problems of type (2), the use of an extended proximal point method is a new idea in the theory of optimization on Riemannian manifolds.

The proximal point algorithm for finding zeros of monotone operators T on Hilbert spaces, see [12], generates a sequence of points $\{p_k\}$ as follows: p_{k+1} is the unique zero of the regularized operator $T+\lambda_k I$, where λ_k is a real number satisfying $0<\lambda_k\leqslant\tilde{\lambda}$, for some $\tilde{\lambda}>0$, and I is the identity operator. The idea is to solve the possibly ill-posed problem of finding zeros of T, by solving a sequence of well-posed problems (i.e., to have exactly one solution when T is strongly monotone) of finding the zeros of $T+\lambda_k I$.

An extension of this problem is the following variational inequality problem: given a convex constraint set C and the monotone operator T find p_* in C such that $\langle T(p_*), p-p_*\rangle\geqslant 0$ for all $p\in C$. When the constraint set of the variational inequality problem is a Riemannian manifold and the operator is a monotone vector field with respect to the metric of the Riemannian manifold, the variational inequality problem becomes the problem of finding the singularities of the monotone vector field.

In the case of Hadamard manifolds we shall solve this problem by extending the proximal point algorithm as follows: We shall generate a sequence $\{p_k\}$, where p_{k+1} is defined as the unique singularity of the regularized vector field $X + \lambda_k \operatorname{grad} \rho_{p_k}$, the sequence $\{\lambda_k\}$ is such that $0 < \lambda_k < \tilde{\lambda}$ for some $\tilde{\lambda} > 0$, the vector field $\operatorname{grad} \rho_{p_k}$ is the gradient vector field of the map $\rho_{p_k} = \frac{1}{2} d^2(., p_k)$ and d is the Riemannian distance.

2. Basic Concepts

In this section some frequently used notations, basic definitions and the important properties of Riemannian manifolds are presented. They can be found in any introductory book on Riemannian geometry(eg. [3,16]). Throughout the paper, all manifolds are smooth, i.e., C^2, paracompact and connected and all functions and vector fields are smooth. C^2 is sufficient for the smoothness of the functions considered and C^1 for the smoothness of the vector fields considered.

Given a Riemannian manifold M, denote the set of vector fields over M by $\mathfrak{X}(M)$, the tangent space of M by T_pM. The metric in M is denoted by $\langle \cdot, \cdot \rangle$ and the corresponding norm by $\| \cdot \|$. Let ∇ be the Levi-Civita connection associated to $(M, \langle , \rangle)$. The *gradient* of a function f, defined in M, is denoted by grad f and by $Hess\, f$ its *Hessian*. The differential of a vector field X is denoted by $A_X v = \nabla_v X$, where v is a tangent vector. The map $P(c)_t^a : T_{c(a)}M \to T_{c(t)}M$, denotes the *parallel transport* along the curve c from $c(a)$ to $c(t)$ and $P(c^{-1})_t^a : T_{c(t)}M \to T_{c(a)}M$ denotes its inverse. Denote the geodesic equation by $\nabla_{\gamma'} \gamma' = 0$. A Riemannian manifold is *complete* if its geodesics are defined for any real values of t. Hopf-Rinow's theorem asserts that if this is the case, any pair of points in M can be joined by a (not necessarily unique) minimal geodesic segment. Moreover, (M, d) is a complete metric space, and bounded and closed subsets are compact. In the paper, all manifolds are assumed to be complete. The *exponential map* $exp_p : T_pM \to M$ is defined by $exp_x v = \gamma_v(1, x)$, where $\gamma(\cdot) = \gamma_v(., p)$ is the geodesic defined by its position p and velocity v at p. A complete, simply connected Riemannian manifold of nonpositive sectional curvature is called a *Hadamard manifold*. From now on, let H be a Hadamard manifold.

Let $\Delta(p_1 p_2 p_3)$ be a geodesic triangle in H and γ_i the geodesic segments joining p_{i+1} and p_{i+2}, where $i = 1, 2, 3\,(mod\,3)$. Set $\ell_i = l(\gamma_i)$ be the length of γ_i and the angles $\theta_i = \sphericalangle\left(\gamma'_{i-1}(0), -\gamma'_{i+1}(l_{i+1})\right)$. From [16] Proposition 4.5, page 223, we obtain

$$\ell_{i+1}^2 + \ell_{i+2}^2 - 2\ell_{i+1}\ell_{i+2}\cos\theta_i \leqslant \ell_i^2. \tag{3}$$

Let M and N be connected Riemannian manifolds and $\Phi: M \to N$ an isometry, i.e., Φ is C^∞, and for all $p \in M$ and $u, v \in T_pM$, we have $\langle d\Phi_p u, d\Phi_p v\rangle = \langle u, v\rangle$. One can verify that Φ preserves geodesics, i.e., β is a geodesic in M if and only if $\gamma = \Phi \circ \beta$ is a geodesic in N, and that $d\Phi_{\gamma(t)}\gamma'(t) = \beta'(t)$. Furthermore, Φ preserves the distance function, i.e., $d(\Phi(p), \Phi(q)) = d(p, q)$, for all $p, q \in M$.

Németh [7], introduced the notion of *monotone vector fields* on M as follows: X is monotone if $\varphi_{(X,\gamma)}(t) = \langle X(\gamma(t)), \gamma'(t)\rangle$ is monotone nondecreasing for all geodesics γ in M. In [4], a vector field X on M was called *strongly monotone* if $\Psi_{(X,\gamma)}(t) = \varphi_{(X,\gamma)}(t) - \lambda\|\gamma'(0)\|^2 t$, is a monotone nondecreasing function of t for some $\lambda > 0$ and all geodesics γ in M. In the case of $M = H$, it has been proved [4] that X is strongly monotone if ond anly if for all p, $q \in H$ it holds that

$$\langle P(\gamma^{-1})_1^0 X(q) - X(p), exp_p{}^{-1}q\rangle \geqslant \lambda d^2(p, q), \tag{4}$$

where $\gamma : [0, 1] \to H$ is the geodesic joining p and q and P is the parallel transport. Furthermore, there exists a unique $\hat{p} \in H$ such that $X(\hat{p}) = 0$.

EXAMPLE 2.1. Take $p' \in H$. The function $\rho_{p'}: H \to R$, defined by

$$\rho_{p'}(p) = \tfrac{1}{2} d^2(p, p'), \tag{5}$$

is smooth and its gradient can be calculated by the formula [16]

$$\operatorname{grad} \rho_{p'}(p) = -\exp_p^{-1} p'. \tag{6}$$

It has been proved [4] that, for all fixed $p' \in H$, the gradient vector field $\operatorname{grad} \rho_{p'}(p)$ is strongly monotone.

EXAMPLE 2.2. A function $f: M \to \mathbb{R}$ is called convex, strictly convex or strongly convex if its composition with each geodesic γ in M is a convex, strictly convex or strongly convex function, respectively. In Ref. [15, 17] it was proved that if f is convex (strictly convex), $\operatorname{grad} f$ is a monotone (strictly monotone) vector field. In Ref. [4] it was proved that if f is strongly convex, $\operatorname{grad} f$ is a strongly monotone vector field.

Examples of monotone vector fields which are of nongradient type can be found in [7–9].

PROPOSITION 1. *Let M, N be Riemannian manifolds and $\Phi: M \to N$ an isometry. The function $f: N \to \mathbb{R}$ is convex iff $g: M \to \mathbb{R}$, defined by $g(p) = f(\Phi(p))$, is convex.*

Proof. It follows from the definition of convexity and the fact that isometries preserve geodesics. □

PROPOSITION 2. *Let M and N be Riemannian manifolds, $X \in \mathfrak{X}(M)$ and $\Phi: M \to N$ an isometry. Let $Y \in \mathfrak{X}(N)$ be defined by $Y = d\Phi \circ X \circ \Phi^{-1}$. Then,*

(i) *X is monotone if and only if Y is monotone;*
(ii) *X is strictly monotone if and only if Y is strictly monotone and*
(iii) *X is strongly monotone if and only if Y is strongly monotone.*

Proof. We shall prove (iii). The proofs of (i) and (ii) are similar. Since Φ is an isometry, $\beta = \Phi^{-1} \circ \gamma$ is a geodesic in M if and only if γ is a geodesic in N and it holds that $\|\gamma'(t)\| = \|\beta'(t)\|$. Then, for all λ, we have

$$\begin{aligned}
\Psi_{(Y,\gamma)}(t) &= \varphi_{(Y,\gamma)}(t) - \lambda \|\gamma'(0)\|^2 t = \langle Y(\gamma(t)), \gamma'(t)\rangle - \lambda\|\gamma'(0)\|^2 t \\
&= \langle d\Phi_{\Phi^{-1}(\gamma(t))}.X(\Phi^{-1}(\gamma(t)))\,,\, \gamma'(t)\rangle - \lambda\|\gamma'(0)\|^2 t \\
&= \langle d\Phi_{\beta(t)} X(\beta(t))\,,\, d\Phi_{\beta(t)}\beta'(t)\rangle - \lambda\|\beta'(0)\|^2 t \\
&= \langle X(\beta(t))\,,\, \beta'(t)\rangle - \lambda\|\beta'(0)\|^2 t = \Psi_{(X,\beta)}(t).
\end{aligned}$$

Therefore, $\Psi_{(Y,\gamma)}$ is monotone for some λ iff $\Psi_{(X,\beta)}$ is monotone. □

3. The Proximal Point Algorithm

3.1. THE PROXIMAL POINT ALGORITHM FOR OPTIMIZATION PROBLEMS

The proximal point algorithm for the minimization of a convex function on a Hadamard manifold was studied in [5]. For a convex function $f: H \to \mathbb{R}$, the proximal point sequence for the minimization of f on H is given by

$$p^{k+1} = \arg\min_{p \in H} \left\{ f(p) + \frac{\lambda_k}{2} d^2(p, p^k) \right\}. \tag{7}$$

We begin by giving some examples of proximal iteration for optimization problem on a Hadamard manifold.

3.1.1. *The space $\mathbb{R}^n$ with other metric*

Endowing $\mathbb{R}^n$ with the metric

$$G(p)=\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & 1 & 0 & 0 \\ 0 & \dots & 0 & 1+4p_{n-1}^2 & -2p_{n-1} \\ 0 & & 0 & -2p_{n-1} & 1 \end{pmatrix},$$

we obtain the Riemannian manifold M_G. Considering $\mathbb{R}^n$ with the usual Euclidean metric, the map $\Phi:\mathbb{R}^n \to M_G$, defined by $\Phi(x) = (x_1, x_2, \dots, x_{n-1}, x_{n-1}^2 - x_n)$ is an isometry. Then, the Riemannian distance in M_G is given by $d^2(p,q) = \|\Phi^{-1}(p) - \Phi^{-1}(q)\|^2 = \sum_{i=1}^{n-1} (p_i - q_i)^2 + (p_{n-1}^2 - p_n - q_{n-1}^2 + q_n)^2$ and the proximal point iteration (7) becomes

$$p^{k+1} = \arg\min_{p\in\mathbb{R}^n} \Big\{ f(p) + \frac{\lambda_k}{2} \sum_{i=1}^{n-1} (p_i - (p^k)_i)^2 + (p_{n-1}^2 - p_n - (p^k)_{n-1}^2 + (p^k)_n)^2 \Big\}.$$

3.1.2. *The positive Orthant with other metric*

Endowing $\mathbb{R}^n$ with the Euclidean metric and $\mathbb{R}^n_{++}$ with the metric $G : \mathbb{R}^n_{++} \to S^n_{++}$,

$$G(p) = \operatorname{diag}\left(p_1{}^{-2}, p_2{}^{-2}, \dots, p_n{}^{-2}\right), \tag{8}$$

we obtain that the mapping $\Phi: \mathbb{R}^n \to \mathbb{R}^n_{++}$,

$$\Phi(x) = (e^{x_1}, e^{x_2}, \dots, e^{x_n}) \tag{9}$$

is an isometry. Then, $d^2(p,q) = \|\Phi^{-1}(p) - \Phi^{-1}(q)\|^2 = \sum_{i=1}^{n} \ln^2 (p_i/q_i)$ and the proximal point iteration (7) becomes

$$p^{k+1} = \arg\min_{p\in\mathbb{R}^n{}_{++}} \left\{ f(p) + \frac{\lambda_k}{2} \sum_{i=1}^{n} \ln^2 \left[\frac{p_i}{(p^k)_i} \right] \right\}.$$

3.1.3. *Hypercube with other metric*

Set

$$Q^n = \left\{(p_1, \ldots, p_n) \in \mathbb{R}^n : |p_i| < \frac{\pi}{2}, i = 1, 2, \ldots, n\right\}$$

and let $\psi\colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ be the function defined by $\psi(\tau) = \ln(\sec\tau + \tan\tau)$. Endowing the Hypercube Q^n with the Riemannian metric $G\colon Q^n \to S^n_{++}$,

$$G(p) = \operatorname{diag}\left(\sec^2 p_1, \sec^2 p_2, \ldots, \sec^2 p_n\right), \tag{10}$$

and $\mathbb{R}^n$ with the Euclidean metric, the mapping $\Phi : Q^n \to \mathbb{R}^n$,

$$\Phi(p) = \left(\psi(p_1), \ldots, \psi(p_n)\right) \tag{11}$$

is an isometry. Then, $d^2(p,q) = \|\Phi(p) - \Phi(q)\|^2 = \sum_{i=1}^n [\psi(q_i) - \psi(p_i)]^2$ and the proximal point iteration (7) becomes

$$p^{k+1} = \arg\min_{p\in Q^n} \left\{ f(p) + \frac{\lambda_k}{2} \sum_{i=1}^n \left[\psi((p^k)_i) - \psi(p_i)\right]^2 \right\}.$$

3.1.4. *The cone of positive semidefinite matrices S^n_{++} with other metric*

Endowing S^n_{++} with the Riemannian metric defined by $\langle U, V\rangle = \operatorname{tr}\left(VX^{-1}UX^{-1}\right)$, we obtain the Riemannian manifold that is complete of curvature $K \leqslant 0$. The Riemannian distance in the manifold S^n_{++} is given by

$$d^2(X,Y) = \sum_{i=1}^n \ln^2 \lambda_i \left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}\right),$$

where $\lambda(A)$ denotes the eigenvalue set of the symmetric matrix A (see [10]). Therefore, the proximal point iteration (7) becomes

$$X_{k+1} = \arg\min_{X\in S^n_{++}} \left\{ f(X) + \frac{\lambda_k}{2} \sum_{i=1}^n \ln^2 \lambda_i \left(X^{-\frac{1}{2}} X_k X^{-\frac{1}{2}}\right) \right\}.$$

3.2. THE PROXIMAL POINT ALGORITHM FOR SINGULARITY PROBLEMS

Let $X \in \mathfrak{X}(H)$ be a monotone vector field and $\mathcal{O}^* \subset H$ the set of singularities of X. The proximal point algorithm for finding zeros of monotone operators was proposed by T. Rockafellar in [12]. We will extend this algorithm for finding singularities of monotone vector fields.

The *proximal point algorithm* for finding a singularity of a monotone vector field on a Hadamard manifold requires one exogenous constant $\tilde{\lambda} > 0$ and one exogenous sequence $\{\lambda_k\}$, satisfying $0 < \lambda_k < \tilde{\lambda}$, for all k. It is defined as follows: take $p_0 \in H$ and define p_{k+1} as the solution of the following equation

$$(X + \lambda_k \operatorname{grad} \rho_{p_k})(p_{k+1}) = 0, \tag{12}$$

where $\rho_{p'}$ is defined in (5). As $\operatorname{grad} \rho_{p_k}$ is strongly monotone and $\lambda_k > 0$, it follows that $X + \lambda_k \operatorname{grad} \rho_{p_k}$ is strongly monotone. Therefore, there exists a unique $p_{k+1} \in H$ such that $(X + \lambda_k \operatorname{grad} \rho_{p_k})(p_{k+1}) = 0$ and our algorithm is well defined. From now on, we will refer to the sequence $\{p_k\}$ generated by (12) as the *proximal sequence*. Note that by (6), it holds that $\operatorname{grad} \rho_{p_k}(p_{k+1}) = -\exp^{-1}_{p_{k+1}} p_k$. Then, equation (12) is equivalent to

$$\lambda_k \exp^{-1}_{p_{k+1}} p_k = X(p_{k+1}). \tag{13}$$

3.2.1. *Convergence of the proximal sequence*

We begin the convergence proof with an auxiliary result. First, we present the well-known concept of Fejér convergence and its application in our context.

In a complete metric space (M, d), the sequence $\{p_k\} \subset M$ is said to be *Fejér convergent* to the nonempty set $U \subset M$, when

$$d(p_{k+1}, y) \leqslant d(p_k, y) \tag{14}$$

for all $y \in U$ and $k \geqslant 0$.

LEMMA 1. *In a complete metric space (M, d), if $\{p_k\} \subset M$ is Fejér convergent to a nonempty set $U \subset M$, then $\{p_k\}$ is bounded. If, furthermore, a cluster point p of $\{p_k\}$ belongs to U, $\lim_{k\to+\infty} p_k = p$.*

Proof. Take $p \in U$. Inequality (14) implies that $d(p_k, p) \leqslant d(p_0, p)$, for all k. Therefore, $\{p_k\}$ is bounded. Take a subsequence $\{p_{k_j}\}$ of $\{p_k\}$ such that $\lim_{k\to+\infty} p_{k_j} = p$. By (14), the sequence of positive numbers $\{d(p_k, p)\}$ is decreasing and it has a subsequence, namely $\{d(p_{k_j}, p)\}$, which converges to 0. Thus, the whole sequence converges to 0, i.e., $\lim_{k\to+\infty} d(p_k, p) = 0$, implying $\lim_{k\to+\infty} p_k = p$. □

LEMMA 2. *If $X \in \mathfrak{X}(H)$ is monotone and $\{p_k\}$ is the proximal sequence, then*

$$d^2(p_{k+1}, p_k) + d^2(p_{k+1}, q) - \frac{2}{\lambda_k}\langle X(p_{k+1}), \exp^{-1}_{p_{k+1}} q\rangle \leqslant d^2(p_k, q), \tag{15}$$

for all $q \in H$.

Proof. Take $q \in H$. Consider the geodesic triangle $\Delta(qp_k p_{k+1})$. From (3), we have

$$d^2(p_{k+1}, p_k) + d^2(p_{k+1}, q) - 2d(p_{k+1}, p_k)d(p_{k+1}, q)\cos\theta \leqslant d^2(p_k, q),$$

where $\theta = \sphericalangle(\exp^{-1}_{p_{k+1}} p_k, \exp^{-1}_{p_{k+1}} q)$, implying that

$$d^2(p_{k+1}, p_k) + d^2(p_{k+1}, q) - 2\langle \exp^{-1}_{p_{k+1}} p_k, \exp^{-1}_{p_{k+1}} q\rangle \leqslant d^2(p_k, q). \tag{16}$$

The statement of the Lemma follows by using (13) in (16). □

THEOREM 1. *If $X \in \mathfrak{X}(H)$ is monotone, $\{p_k\}$ is the proximal sequence and $\mathcal{O}^*$ is non-empty, then $\lim_{k\to+\infty} p_k = p_*$ for some $p_* \in H$.*

Proof. Take $\tilde{p} \in \mathcal{O}^*$. The monotonicity of X implies $\langle X(p_{k+1}), \exp^{-1}_{p_{k+1}} \tilde{p}\rangle \leqslant \langle X(\tilde{p}), \exp^{-1}_{p_{k+1}} \tilde{p}\rangle$. Since $X(\tilde{p}) = 0$, we have that

$$\langle X(p_{k+1}), \exp^{-1}_{p_{k+1}} \tilde{p}\rangle \leqslant 0. \tag{17}$$

Then, substituting q by $\tilde{p}$ in (15) and by using (17), we get

$$0 \leqslant d^2(p_{k+1}, p_k) \leqslant d^2(p_k, \tilde{p}) - d^2(p_{k+1}, \tilde{p}). \tag{18}$$

The inequality (18) implies that $\{p_k\}$ is Fejér convergent to the set $\mathcal{O}^*$ and that $\lim_{k\to\infty} d^2(p_{k+1}, p_k) = 0$. By Lemma 1, there exists a subsequence $\{p_{k_j}\}$ of $\{p_k\}$ which converges to some $p_* \in H$. It holds that $d(p_{k_j+1}, p_{k_j}) = \|\exp^{-1}_{p_{k_j+1}} p_{k_j}\|$. Then, by (13)

$$\lambda_{k_j} d(p_{k_j+1}, p_{k_j}) = \lambda_{k_j} \|\exp^{-1}_{p_{k_j+1}} p_{k_j}\| = \|X(p_{k_j+1})\|. \tag{19}$$

Since X is continuous, $\{p_{k_j}\}$ is convergent to p_* and $0 < \lambda_k < \tilde{\lambda}$, (18) implies that $\lim_{k\to\infty} d^2(p_{k_j+1}, p_{k_j}) = 0$. Hence, (19) yields,

$$\|X(p_*)\| = \lim_{j\to+\infty} \|X(p_{k_j})\| = \lim_{j\to+\infty} \lambda_{k_j} d(p_{k_j+1}, p_{k_j}) = 0,$$

implying that $p_* \in \mathcal{O}^*$. Therefore, by Lemma 1, $\lim_{k\to\infty} p_k = p_*$ and the proof is complete. □

3.2.2. *Invariance of the proximal sequence through isometries*

Isometric manifolds bear the same properties from Riemannian geometric viewpoint, but on some of them calculus is much easier. On the other hand, in this subsection we shall show that proximal sequences are invariant through isometries. The above remarks will be exploited in the example of Section 3.2.3.

PROPOSITION 3. *Let* H_1, H_2 *be Hadamard manifolds and* $\Phi : H_1 \to H_2$ *an isometry. If* $\{p_k\}$ *is the proximal sequence on* H_1 *with the starting point* $p_0 \in H_1$ *associated to the vector field* $X \in \mathfrak{X}(H_1)$ *and the sequence* $\{\lambda_k\}$*, then* $\{\Phi(p_k)\}$ *is the proximal sequence on* H_2 *with the starting point* $\Phi(p_0)$ *associated to the vector field* $Y = d\Phi \circ X \circ \Phi^{-1} \in \mathfrak{X}(H_2)$ *and the sequence* $\{\lambda_k\}$.

Proof. Since Φ is an isometry, the geodesics of H_1 are transformed into geodesics of H_2 such that the tangent vector of a geodesic on H_1 is transformed into the tangent vector of its transformed geodesic on H_2. Hence, we have

$$d\Phi(p_{k+1})\left(\exp^{-1}_{p_{k+1}} p_k\right) = \exp^{-1}_{\Phi(p_{k+1})} \Phi(p_k). \tag{20}$$

The proximal sequence $\{p_k\}$ on H_1, with respect to a starting point $p_0 \in H_1$, associated to the vector field $X \in \mathfrak{X}(H_1)$ and the sequence $\{\lambda_k\}$ is given by

$$\lambda_k \exp^{-1}_{p_{k+1}} p_k = X\left(p_{k+1}\right). \tag{21}$$

Since $Y \circ \Phi = d\Phi \circ X$, by applying $d\Phi(p_{k+1})$ to (21) and by using (20), we obtain

$$\lambda_k \exp^{-1}_{\Phi(p_{k+1})} \Phi(p_k) = Y(\Phi(p_{k+1})).$$

Hence, $\{\Phi(p_k)\}$ is the proximal sequence on H_2, with respect to the starting point $\{\Phi(p_0)\}$, associated to the vector field $Y \in \mathfrak{X}(H_2)$ and the sequence $\{\lambda_k\}$. □

With the notations of Proposition 3, we have as follows:

COROLLARY 1. *If the proximal sequence* $\{p_k\}$ *is convergent to a singularity* p_* *of* X*, then the proximal sequence* $\{\Phi(p_k)\}$ *is convergent to the singularity* $\Phi(p_*)$ *of* Y.

Proof. It follows immediately from the equality $Y \circ \Phi = d\Phi \circ X$. □

3.2.3. *Example*

Let $\mathbb{H}^n$ be the n dimensional hyperbolic space of constant sectional curvature $K = -1$. Consider the following model for $\mathbb{H}^n$:

$$M = \left\{\xi = (\xi_1, \ldots, \xi_n, \xi_{n+1}) \in \mathbb{R}^{n+1} :\ \xi_{n+1} > 0 \text{ and } \{\xi, \xi\} = -1\right\},$$

where for the vectors $\xi = (\xi_1, \ldots, \xi_{n+1})$, $\eta = (\eta_1, \ldots, \eta_{n+1}) \in \mathbb{R}^{n+1}$ and $\{\xi, \eta\} = \xi_1\eta_1 + \cdots + \xi_n\eta_n - \xi_{n+1}\eta_{n+1}$. The metric of M is induced from the Lorentz metric $\{.,.\}$ of $\mathbb{R}^{n+1}$ and it will be denoted by the same symbol.

Then, a normalized geodesic γ_x of $\mathbb{H}^n$ starting from $x(\gamma_x(0)=x)$, will have the equation

$$\gamma_x(t)=(\cosh t)x+(\sinh t)v, \tag{22}$$

where $v=\dot{\gamma}_x(0)\in T_x\mathbb{H}^n$ is the tangent unit vector of γ in the starting point. We also have $\{u, x\}=0$, for all $u\in T_x\mathbb{H}^n$. Equation (22) implies $\exp tv=(\cosh t)x+(\sinh t)v$, for any unit vector v and

$$\exp_x^{-1} y=\operatorname{arccosh}(-\{x, y\})\frac{y+\{x, y\}x}{\sqrt{\{x, y\}^2-1}}, \tag{23}$$

for all $x, y\in\mathbb{H}^n$ and $v\in T_x\mathbb{H}^n$. This model of the hyperbolic space is called the Minkowski model. Next, consider the following model for $\mathbb{H}^n$: $U=\{x=(x_1, x_2, \dots, x_n)\in\mathbb{R}^n : x_n>0\}$. The set U is the upper half-plane of dimension n. Endowing U with the Riemannian metric defined by matrix $G=(g_{ij})$, where

$$g_{11}(x_1, \dots, x_n)=\cdots=g_{nn}(x_1, \dots, x_n)=\frac{1}{x_n},\quad g_{ij}(x_1, \dots, x_n)=0,\ \text{ if } i\neq j.$$

we obtain the upper half-plane model of the Hyperbolic space $\mathbb{H}^n$.
Consider the case $n=2$. It can be seen that the map $\Phi: M\to U$ given by the equation

$$(x_1, x_2, x_3)\mapsto\frac{2}{x_3-x_2}(x_1, 1), \tag{24}$$

is an isometry between M and U with inverse $\Phi^{-1}: U\to M$ given by the equation

$$(x_1, x_2)\mapsto\frac{1}{4x_2}\left(4x_1, x_1^2+x_2^2-4, x_1^2+x_2^2+4\right).$$

By (23), if X is a vector field on M and $\{\lambda_k\}$ is an exogenous sequence, the proximal sequence $\{p^k\}$, with respect to a starting point $p^0\in M$, X and $\{\lambda_k\}$ is given by the recurrence

$$\operatorname{arccosh}\left(-\left\{p^{k+1}, p^k\right\}\right)\frac{p^k+\left\{p^{k+1}, p^k\right\}p^{k+1}}{\sqrt{\left\{p^{k+1}, p^k\right\}^2-1}}=X\left(p^{k+1}\right).$$

If $Y=d\Phi\circ X\circ\Phi^{-1}$ is the transformed vector field of X on U, with respect to Φ, $\{\Phi(p^k)\}$ is the proximal sequence, with respect to the starting point $\{\Phi(p^0)\}$, to Y and $\{\lambda_k\}$. If X is monotone and has at least one singularity, the proximal sequence $\{p^k\}$ is convergent to a singularity p^* of X. In this case, the proximal sequence $\{\Phi(p^k)\}$ is convergent to the singularity $\Phi(p^*)$ of the monotone vector field Y.

In [7] it is shown that the vector field $X(x_1, x_2, x_3) = (x_1x_3, x_2x_3, x_3^2 - 1)$ on M is strictly monotone. The only singularity of X is $(0, 0, 1)$. The proximal sequence $\{p^k\}$, with respect to a starting point $p^0 \in M$, X and $\{\lambda_k\}$ is given by the recurrence

$$\operatorname{arccosh}\left(-\left\{p^{k+1}, p^k\right\}\right) \frac{p^k + \left\{p^{k+1}, p^k\right\} p^{k+1}}{\sqrt{\left\{p^{k+1}, p^k\right\}^2 - 1}} = \left(p_1^{k+1} p_3^{k+1}, p_2^{k+1} p_3^{k+1}, \left(p_3^{k+1}\right)^2 - 1\right),$$

and is convergent to $(0, 0, 1)$. It is easy to calculate that the image of X through Φ is $Y = \frac{1}{32}\left(16x_1x_2, 2x_1^2x_2^2 - 8x_1^2 + 8x_2^2 - 32\right)$ and $\Phi(0, 0, 1) = (0, 2)$. Y is strictly monotone on U. By (24) $\{(2/(p_3^k - p_2^k))(p_1^k, 1)\}$ is the proximal sequence, with respect to the starting point $(2/(p_3^0 - p_2^0))(p_1^0, 1)$, Y and $\{\lambda_k\}$. It is convergent to $(0, 2)$ the only singularity of Y.

4. Problems From the Geometric Viewpoint

Let us consider again the problems (1) and (2) which are nonconvex and nonmonotone in the original representation. Choosing an appropriate metric, the problems can be transformed into convex and monotone ones, respectively. We note that our goal is to transform nonconvex (nonmonotone) problems into Hadamard convex (Hadamard monotone) ones. In this viewpoint, we are not interested in whether the curvature is zero or not, only its sign is important.

4.1. THE PLANE WITH OTHER METRICS

Consider the following unconstrained problems defined in the Euclidean plane.

PROBLEM 4.1. *In optimization problem* (1), *take the Rosenbock's banana function* $f: \mathbb{R}^2 \to \mathbb{R}$, $f(p_1, p_2) = 100\left(p_2 - p_1{}^2\right)^2 + \left(1 - p_1\right)^2$.

PROBLEM 4.2. *In problem* (2), *take* $X: \mathbb{R}^2 \to \mathbb{R}^2$, $X(p) = (-p_1^2 + p_1 + p_2, -2p_1^3 + 2p_1^2 + 2p_1p_2 - p_1)$.

Problem 4.1 is not convex in the classical sense, i.e., the objective function f is not convex, and Problem 4.2 is not monotone in the classical sense, i.e., the vector field X is not monotone. Endowing $\mathbb{R}^2$ with the Riemannian metric $G:\mathbb{R}^2 \to S^n_{++}$,

$$G(p_1, p_2) = \begin{pmatrix} 1+4p_1^2 & -2p_1 \\ -2p_1 & 1 \end{pmatrix},$$

we obtain the Riemannian manifold M_G that is complete and of constant curvature $K=0$. Note that the map $\Phi:\mathbb{R}^2 \to M_G$, $\Phi(x_1, x_2) = \left(x_1, x_1^2 - x_2\right)$ is an isometry. Now, define the convex function $g:\mathbb{R}^2 \to \mathbb{R}$ by $g(x_1, x_2) = 100x_2^2 + (1-x_1)^2$ and observe that $g(x_1, x_2) = f(\Phi(x_1, x_2))$. Therefore, by Proposition 1, it follows that f is convex in M_G. Let $Y:\mathbb{R}^2 \to \mathbb{R}^2$ be a monotone vector field defined by $Y(x_1, x_2) = (x_1 - x_2, x_1)$. Note that $X = d\Phi \circ Y \circ \Phi^{-1}$. Therefore, by Proposition 2, X is monotone in M_G.

PROBLEM 4.3. *In problem* (1), *take* $f:\mathbb{R}^2 \to \mathbb{R}$, $f(p_1, p_2) = e^{p_1}\left(\cosh(p_2) - 1\right)$.

PROBLEM 4.4. *In problem* (2), *take* $X:\mathbb{R}^2 \to \mathbb{R}^2$, $X(p_1, p_2) = \left(e^{p_1}(\cosh(p_2) - 1)\ e^{-p_1}\sinh(p_2)\right)$.

Problem 4.3 is not convex in the classical sense, i.e., the objective function f is not convex, and Problem 4.4 is not monotone in the classical sense, i.e., the vector field X is not monotone. Endowing $\mathbb{R}^2$ with the Riemannian metric $G:\mathbb{R}^2 \to S^n_{++}$,

$$G(p_1, p_2) = \begin{pmatrix} 1 & 0 \\ 0 & e^{2p_1} \end{pmatrix},$$

we obtain the Riemannian manifold M_G that is complete and of constant curvature $K=-1$. The Christoffel symbols are given by

$$\Gamma_{11}^1 = \Gamma_{11}^2 = \Gamma_{12}^1 = \Gamma_{21}^1 = \Gamma_{22}^2 = 0,\ \Gamma_{12}^2 = \Gamma_{21}^2 = 1 \text{ and } \Gamma_{22}^1 = -e^{2p_1}.$$

Then, for each vector field $Y(p_1, p_2) = \left(a(p_1, p_2), b(p_1, p_2)\right)$, defined on M_G, we have

$$A_Y(p_1, p_2) = \begin{pmatrix} \frac{\partial a}{\partial p_1} & \frac{\partial a}{\partial p_2} - e^{2p_1} b \\ e^{2p_1}\left(\frac{\partial b}{\partial p_1} + b\right) & e^{2p_1}\left(\frac{\partial b}{\partial p_2} + a\right) \end{pmatrix}. \tag{25}$$

The gradient vector field of f is grad $f(p)=G^{-1}(p)(\partial f/\partial p_1(p), \partial f/\partial p_2(p))$. From (25), it follows that the Hessian matrix $\mathrm{Hess}(f)=A_{\mathrm{grad}(f)}$ is given by

$$\mathrm{Hess} f(p_1, p_2)=\begin{pmatrix} e^{p_1}(\cosh(p_2)-1) & 0 \\ 0 & e^{p_1}\cosh(p_2)+e^{3p_1}(\cosh(p_2)-1)\end{pmatrix}.$$

Note that this matrix is positive semidefinite. Therefore, f is convex in M_G. It can also be checked that

$$A_X(p_1, p_2)=\begin{pmatrix} e^{p_1}(\cosh(p_2)-1) & 0 \\ 0 & e^{p_1}\cosh(p_2)+e^{3p_1}(\cosh(p_2)-1)\end{pmatrix}.$$

Thus, X is monotone in M_G, see Ref. [4].

4.2. THE POSITIVE ORTHANT WITH OTHER METRICS

Consider the following constrained problems defined in the positive orthant.

PROBLEM 4.5. *In optimization problem* (1)*, take the poseynomial* $f:\mathbb{R}^n_{++}\to\mathbb{R}$,

$$f(p_1, \dots, p_n)=\sum_{i=1}^{m} c_i \prod_{j=1}^{n} p_j^{b_{ij}},$$

where $c_i\in\mathbb{R}_{++}$ *and* $b_{ij}\in\mathbb{R}$ *for all* i, j.

PROBLEM 4.6. *In problem* (2)*, take the vector field* $X:\mathbb{R}^n_{++}\to\mathbb{R}^n$*, defined by* $X(p_1, \dots, p_n)=(a_1, \dots, a_n)$*, where* $a_i=p_i\ln(p_1\dots p_i p_{i+1}^{-1}\dots p_n^{-1})$ *for all* $i=1, \dots, n$.

Problem 4.5 is not convex in the classical sense, i.e., the objective function f is not convex, and Problem 4.6 is not monotone in the classical sense, i.e., the vector field X not monotone. Endowing $\mathbb{R}^n_{++}$ with the Riemannian metric G, defined in (8), we obtain the Riemannian manifold M_G that is complete and of constant curvature $K=0$. Now, define the function $g:\mathbb{R}^n\to\mathbb{R}$ by

$$g(x_1, \dots, x_n)=\sum_{i=1}^{m} c_i e^{\sum_{j=1}^{n} b_{ij}x_j}.$$

Note that g is convex in the classical sense and that $g(x_1, \dots, x_n)=f(\Phi(x_1, \dots, x_n))$, where Φ is the isometry defined in (8). Therefore, by

Proposition, 1, it follows that f is convex in M_G. Let $Y:\mathbb{R}^n \to \mathbb{R}^n$ be the monotone vector field, defined by $Y(x)=Ax$, where $x=(x_1,\dots,x_n)$ and

$$A=\begin{pmatrix} 1 & -1 & \dots & -1 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ 1 & \dots & 1 & 1 \end{pmatrix}, \qquad (26)$$

Note that $Y=d\Phi\circ X\circ\Phi^{-1}$, where Φ is the isometry defined in (8). Hence, from Proposition 2, X is monotone in M_G.

4.3. THE HYPERCUBE WITH OTHER METRIC

Consider the following problems:

PROBLEM 4.7. *In problem* (1), *take* $f: Q^n \to \mathbb{R}$, *defined by* $f(p_1,\dots,p_n) = \psi(p_1)+\dots+\psi(p_n)$.

PROBLEM 4.8. *In problem* (2), *take* $X: Q^n \to \mathbb{R}^n$, *defined by* $X(p_1,\dots,p_n)= (a_1,\dots,a_n)$, *where* $a_i = \cos(p_i)\left(\sum_{j\leqslant i}\psi(p_j)-\sum_{j>i}\psi(p_j)\right)$, *for all* $i=1,\dots,n$.

Problem 4.7 is not convex in the classical sense, i.e., the objective functions f is not convex, and Problem 4.8 is not monotone in the classical sense, i.e., the vector field X is not monotone. Endowing Q^n with the Riemannian metric G defined in (10), we obtain the Riemannian manifold M_G that is complete and of constant curvature $K=0$. Now, define a convex function $g:\mathbb{R}^n\to\mathbb{R}$ by $g(x_1,\dots,x_n)=x_1+\dots+x_n$ and observe that $f(p_1,\dots,p_n)=g(\Phi(p_1,\dots,p_n))$, where g is an isometry defined in (11). Therefore, by Proposition 1, it follows that f is convex in M_G. Let $Y:\mathbb{R}^n\to\mathbb{R}^n$ be defined by $Y(x)=Ax$, where A is the matrix (26). Taking Φ, the isometry defined in (11), we obtain that $X=d\Phi^{-1}\circ Y\circ\Phi$. Hence, by Proposition 2, X is monotone in M_G.

4.4. THE CONE OF THE POSITIVE SEMIDEFINITE MATRICES WITH OTHER METRIC

Consider the following constraint problems:

PROBLEM 4.9. *In optimization problem* (1), *take* $f(X)=\left(\ln\det X\right)^2$ *and* $M=S^n_{++}$.

PROBLEM 4.10. *In problem* (2), *take* $T(X)=2\left(\ln\det X\right)X$ *and* $M=S^n_{++}$.

Problem 4.9 is not convex in the classical sense, i.e., the objective function f is not convex, and Problem 4.10 is not monotone in the classical sense, i.e., the vector field X is not monotone. Endowing S^n_{++} with the Riemannian metric defined in Section 3.1.4, its geodesic equation becomes

$$\zeta''(t) = \zeta'(t)\zeta^{-1}(t)\zeta'(t), \tag{27}$$

see [10]. A function f, defined on S^n_{++}, is convex if and only if for any geodesic ζ in S^n_{++}

$$\text{Hess}\, f_{\zeta(t)}\big(\zeta'(t), \zeta'(t)\big) = \text{tr}\big(f''(\zeta(t))\zeta'(t), \zeta'(t)\big) + \text{tr}\big(f'(\zeta(t)), \zeta''(t)\big) \geqslant 0, \tag{28}$$

that is, the Hessian matrix of the function f is positive semidefinite. Therefore, from equations (27), (28) and the definition of the Hessian, it follows that function f is convex in S^n_{++} if it satisfies the condition

$$\text{tr}\big(Vf''(X)V\big) + \text{tr}\big(VX^{-1}Vf'(X)\big) \geqslant 0, \tag{29}$$

for all $X \in S^n_{++}$ and $V \in S^n$. It can be checked whether f satisfies the condition (29) and grad $f(X) = T(X)$. Hence, f is convex and T is monotone (see Example 2.2).

5. Final Remark

Here we presented a novel method of finding the singularities of monotone vector fields on Hadamard manifolds by using an extension of the classical proximal point method of Rockafellar for finding zeros of monotone operators. Whether Rockafellar's method can be extended to more general Riemannian mamifolds or not is still unclear.

Acknowledgements

The authors would like to thank the two anonymous referees for their constructive comments which have helped to substantially improve the presentation of the paper. The first author was supported in part by PRONEX-Otimização(FAPERJ/CNPq) The second author was supported in part by CAPES, FUNAPE/UFG, PADCT-CNPq and PRONEX-Otimização (FAPERJ/CNPq). The third author was supported in part by PADCT-CNPq and PRONEX-Otimização(FAPERJ/CNPq). The fourth author was supported in part by grant No.T029572 of the National Research Foundation of Hungary and by the Hungarian Research Grant OTKA T043276.

References

1. Avriel, M. (1976), *Nonlinear Programming*, Prentice-Hall, Inc. Englewood Cliffs, NJ.
2. Burachik, R.S., Graña Drummond, L.M., Iusem, A.N. and Svaiter, B.F. (1995), Full convergence of the steepest descent method with inexact line searches, *Optimization* 32, 137–146.
3. do Carmo, M.P. (1992), *Riemannian Geometry*, Birkhaüser, Boston.
4. da Cruz Neto, J.X., Ferreira, O.P. and Lucambio Pérez, L.R. (2002), Contribution to the study of monotone vector fields, *Acta Mathematica Hungarica* 94(4), 307–320.
5. Ferreira, O.P. and Oliveira, P.R. (2002), Proximal point algorithm on Riemannian manifold, *Optimization* 51(2), 257–270.
6. Iusem, A.N. and Svaiter, B.F. (1995), A proximal regularization of the steepest descent method, *Recherche opérationnelle/Operations Research* 29(2), 123–130.
7. Németh, S.Z. (1999), Geodesic monotone vector fields, *Lobachevskii Journal of Mathematics* 5, 13–28.
8. Németh, S.Z. (1999), Monotone vector fields, *Publicationes Mathematicae Debrecen*, 54(3–4), 437–449.
9. Németh, S.Z. (1999), Monotonicity of the complementarity vector field of a nonexpansive map, *Acta Mathematica Hungarica* 84(3), 189–197.
10. Nesterov, Y.E. and Todd, M.J. (2002), On the Riemannian geometry defined by self-concordant barriers and interior-point methods, *Found. Comp. Math.* 2(4), 333–361.
11. Ortega, J.M. and Rheimboldt, W.C. (1970), *Interactive Solution of Nonlinear Equations in Several Variables*, Academic Press, New York.
12. Rockafellar, R.T. (1976), Monotone operators and the proximal point algorithm, *SIAM Journal of Control and Optimization*, 14, 877–898.
13. Rapcsák, T. (1996), Geodesic convexity on $R^n{}_{++}$, *Optimization* 37, 341–355.
14. Rapcsák, T. (1998), Variable metric methods along geodesics, In: Giannessi, F., Komlósi, S. and Rapcsak, T. (eds), *New trends in Mathematical Programming*, Kluwer Academic Publishers, Dordrecht.
15. Rapcsák, T. (1997), *Smooth Nonlinear Optimization in* R^n, Kluwer Academic Publishers, Dordrecht.
16. Sakai, T. (1996), *Riemannian Geometry*, Translations of Mathematical Monographs vol. 149, American Mathematical Society, Providence, R.I.
17. Udrişte, C. (1994), *Convex Functions and Optimization Methods on Riemannian Manifolds*, Mathematics and its Applications vol. 297, Kluwer Academic Publishers, Dordrecht.
18. Udrişte, C. (1996), Riemannian convexity in programming (II), *Balkan Journal of Geometry and Its applications*, 1(1), 99–109.
19. Udrişte, C. (1996), Sufficient decrease principle on Riemannian manifolds, *Balkan Journal of Geometry and Its applications*, 1(2), 111–123.
20. Tsagas, Gr. and Udrişte, C. (2002), *Vector Fields and Their Applications*, Geometry Balkan Press, Bucharest.