NEW RESULTS ON LINEAR OPTIMIZATION THROUGH DIAGONAL METRICS AND RIEMANNIAN GEOMETRY TOOLS

E. A. Papa Quiroz and P. Roberto Oliveira

Programa de Engenharia de Sistemas e Computação -COPPE
Federal University of Rio de Janeiro
Rio de Janeiro, Brazil
erik@cos.ufrj.br, poliveir@cos.ufrj.br

May 2004

Abstract

The Riemannian Geometry has been utilized in Mathematical Optimization as an important tool for the analysis of continuous trajectory, convergence and complexity of algorithms, as well to obtain new class of methods. In this paper, we generalize the invariant Riemannian metric concept and study the general diagonal metrics defined on the manifolds $I\!R^n_{++}$ and $\langle 0,1\rangle^n$, obtaining some geometric properties useful to the development of new interior point methods.

We propose the diagonal Riemannian metrics $\pi^2 \csc^4(\pi x)$ and $X^{-2}(I-X)^{-2}$ on the manifold $\langle 0,1\rangle^n$. We get two new examples of geodesic algorithms and a proximal point algorithm to solve the problem min $f(x)$ s.to: $0 \leq x_i \leq 1,\ \ i=1,2,...,n$, whose convergence results are known, for general metrics. Finally, using those metrics we introduce two barrier functions:

$$b_1(x) = -\frac{1}{6}\sum_{i=1}^{n}\{4\ln(\sin \pi x_i) - \cot^2(\pi x_i)\},$$

$$b_2(x) = \sum_{i=1}^{n}(2x_i - 1)[\ln x_i - \ln(1-x_i)].$$

We show that those functions have self-concordant properties. The first barrier is a self-concordant function and the second one is a $(9/4)n$-self-concordant barrier. Thus, in the last case, we have a polynomial algorithm to solve the problem min $c^T x$ s.to: $Ax = b$; $0 \leq x_i \leq 1,\ \ i = 1,2,...,n$, with similar complexity to classic logarithm barrier applied to the first octant.

Keywords: Riemannian metric, geodesic algorithm, self-concordant function, barrier function, interior point methods.

1 Introduction

The Linear Optimization (LO in the sequel) has a broad range of applications in the most diverse fields of sciences and engineering. For instance, in economics, finances, industries, telecommunications, transports, etc. Due to that, LO is object of constant study by the researchers that act on those areas.

Since Karmarkar [7] introduced his polynomial-time algorithm for LO, the field of interior point methods for both Linear Optimization and certain Convex Optimization problems have been developed at a rapid rate, due to their excellent computational and theoretical properties; see for example, den Hertog [4], Nesterov and Nemirovski [11], S.Wright [21]. On the other hand, the perspective of metrics that underlies continuous optimization is evident in many algorithm developments and their theoretical analysis, see Nazareth, [10], for nonlinear unconstrained minimization; for interior point methods, see Karmarkar, [8], and Cruz Neto and Oliveira, [1]. An interesting point is the exploring of the dependent metric gradient concept, that allows the construction of a large set of directions, including some provided by interior point methods, which can be seen as Cauchy (or gradient) ones, see the quoted [8] and [1]. Then, as Luenberger [9] did with his descent geodesic method, but applying the more general framework of Riemannian manifolds, it is possible to show the convergence and linear rate of convergence for a large class of primal algorithms applied to linear and nonlinear convex optimization problems, see Cruz Neto et, al. [2]. In the same way, Ferreira and Oliveira [5], [6] generalized, respectively, the sub gradient and proximal point methods to the context of Riemannian manifolds. Karmarkar [8], using tools of Riemannian Geometry, proved that the complexity of his algorithm is related to the curvature of the trajectory. The deep analysis of Nesterov and Todd [12] on the Riemannian Geometry defined by self-concordant barriers, leads the authors to interesting conclusions about the optimality of the trajectories of primal-dual polynomial algorithms that are near to some geodesic.

The diagonal class of metrics $(\mathbb{R}^n_{++}, X^{-r})$, for $r \geq 1$ has been utilized to create new families of interior point algorithms, see Saigal [19], a primal algorithm applied to LO, Pinto et al.[18], a primal method for convex linear problems, Den Hertog, [4], in a barrier context, Oliveira and Oliveira [13], a proximal algorithm for convex problems in the non negative octant, and Pereira and Oliveira [17], which considered the nonlinear complementarity problem. In this paper, we make some advances on those ideas. First, we observe that, in the point of view of the application of Riemannian geometry tools, the minimum to demand is the completeness of the manifold, this means that there is a geodesic between any two points, so we can measure distances in appropriate way, through the minimum length geodesic. Second, in the point of view of the interior point theory, self-concordant barrier are welcome. Both properties are, in general, absent on the quoted papers, and we will be able to ensure part of them.

The paper is organized as follows. In the next section we review some basic facts on Riemannian Geometry. In section 3 we generalize the method to obtain invariant Riemannian metric and study the general diagonal metrics on the feasible regions of the linear optimization problems, considered as Riemannian manifolds, specifically on the positive octant $\mathbb{R}^n_{++}$, the hypercube $C^n_0 = \langle 0, 1\rangle^n$ and the product of positive octants $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$. In section 4, we study the geometric properties of the general diagonal metric, obtaining simple equations for geodesic curves in closed form, null curvature of the Riemannian manifolds, sufficient conditions that ensure completeness, explicit expressions for gradient and Hessian of smooth functions, explicit geodesic curves for certain Riemannian manifolds defined on $\mathbb{R}^n_{++}$ and $C^n_0 = \langle 0, 1\rangle^n$. In section 5, we present some applications. First, introducing the metrics $\pi^2 \csc^4(\pi x) = diag(\pi^2 \csc^4(\pi x_1), \pi^2 \csc^4(\pi x_2), ..., \pi^2 \csc^4(\pi x_n))$ and $X^{-2}(I - X)^{-2}$

on the hypercube, we present new examples of geodesic algorithms to solve $\min f(x)$ s.to: $0 \leq x_i \leq 1,\ \forall i = 1, 2, ..., n$, and a proximal point algorithm for the same problem. Second, introducing the barrier functions

$$b_1(x) = -\frac{1}{6\pi^2}\sum_{i=1}^{n}\{4\ln(\sin \pi x_i) - \cot^2(\pi x_i)\}$$

$$b_2(x) = \sum_{i=1}^{n}(2x_i - 1)[\ln x_i - \ln(1 - x_i)].$$

we show that they are self-concordant, and the second is also a (9/4)n-self-concordant barrier. Thus we have a new polynomial algorithm for the problem $\min c^T x$ s.to: $Ax = b;\quad 0 \leq x_i \leq 1, \quad i = 1, 2, ..., n$ with similar complexity to the classic logarithm barrier applied to the hypercube.

2 Some Tools of Riemannian Geometry

In this section we introduce some fundamental properties and notation on Riemannian manifolds. Those basic facts can be seen, for example, in do Carmo [3].

Let S be a differential manifold. We denote by T_xS the tangent space of S at x and $TS = \bigcup_{x \in S} T_xS$. T_xS is a linear space and has the same dimension of S. Because we restrict ourselves to real manifolds, T_xS is isomorphic to $I\!R^n$. If S is endowed with a Riemannian metric g, then S is a Riemannian manifold and we denoted it by (S, g). Observe that g can always be represented by some matrix. The inner product of two vectors $u, v \in T_xS$ is written as $\langle u, v\rangle_x := g_x(u, v)$, where g_x is the metric evaluated at the point x. The norm of a vector $v \in T_xS$ is $||v|| := \sqrt{\langle v, v\rangle_x}$. A Lie Group is a differential manifold with a group structure $\bullet$ such that the map $\rho : S \times S \to S$ with $\rho(x, y) = x \bullet y^{-1}$ is differentiable. For an element $y \in S$ the left translation by x is the map $L_x : S \to S$ defined by $L_x(y) = x \bullet y$. Note that the left translation L_x is a differentiable difeomorphism. Let S be a Lie group, a Riemannian metric g on S is said to be *left invariant* if for each $x \in S$ the left translation L_x is an isometry, that is

$$< u, v >_y = < d(L_x)_y u, d(L_x)_y v >_{L_x y},\ \forall y \in S. \tag{2.1}$$

The metric can be used to define the length of a piecewise smooth curve $c : [a, b] \to S$ joining x' to x by $L(c) = \int_a^b \|c'(t)\|dt$, where $c(a) = x'$ and $c(b) = x$. Minimizing this length functional over the set of all curves we obtain a Riemannian distance $d(x', x)$ which induces the original topology on S.

Given two vector fields $X, Y : S \to TS$, along a smooth curve $\alpha : [a, b] \to S$, the covariant derivative of Y in the direction X is $\nabla_X Y$. Levi-Civita theorem ensures that, given a Riemannian manifold (S, g), there exists a unique connection ∇, which defines the covariant derivative D/dt, is symmetric and compatible with the metric (that means: $d/dt < X, Y > = < DX/dt, Y > + < X, DY/dt >$. That is called a Riemannian connection. A curve $\alpha(t)$ is a geodesic starting from a point p with direction v, if $\alpha(0) = p$, $\alpha'(0) = v$ and

$$\frac{d^2\alpha_k}{dt^2} + \sum_{i,j}\Gamma^k_{ij}\frac{d\alpha_i}{dt}\frac{d\alpha_j}{dt} = 0, \quad k = 1, ..., n \tag{2.2}$$

where Γ^k_{ij} are the Christoffel symbols, expressed by

$$\Gamma^m_{ij} = \frac{1}{2}\sum_k\left\{\frac{\partial}{\partial x_i}g_{jk} + \frac{\partial}{\partial x_j}g_{ki} - \frac{\partial}{\partial x_k}g_{ij}\right\}g^{km}, \tag{2.3}$$

(g^{ij}) denoting the inverse matrix of the metric $g = (g_{ij})$. Now, suppose that X and Y are represented by $X = \sum_{i=1}^{n} u^i X_i$, $Y = \sum_{i=1}^{n} v^i X_i$, for some local basis $\{X_i\}$ for $T_x S$, then $\nabla_{X_i} X_k = \sum_{j=1}^{n} \Gamma_{ik}^j X_j$. A Riemannian manifold is complete if its geodesics are defined for any value of t. We denote by R the curvature tensor defined by

$$R(X,Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[X,Y]} Z \tag{2.4}$$

where X, Y, Z are vector fields of S and $[X,Y] := YX - XY$ is the Lie Bracket. Indeed, that formula can be simplified, as $[X_i, X_j] = 0$. Clearly, the curvature, in a Riemannian manifold, is tri-linear. Now, the sectional curvature with respect to X and Y is defined by

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

If $K(X,Y) = 0$, S is a null curvature Riemannian manifold.

The gradient of a differentiable function $f : S \to I\!R$, $\mathtt{grad}f$, is a vector field on S defined by $df(X) = \langle \mathtt{grad}f, X\rangle = X(f)$, where X is also a vector field on S. Now, if f is twice-differentiable we can define the Hessian H^f, as the covariant derivative of the gradient vector field, that is, $H^f = D\mathtt{grad}\, f/dt$. Thus, the Hessian of f at a point $x \in S$ on the direction $v \in T_x S$ is $H_x^f(v) = (D\mathtt{grad}\, f/dt)(x) = \nabla_v \mathtt{grad}\, f\,(x)$.

3 Invariant Riemannian Metrics

Karmarkar [8] used the invariance under translation property of the Riemannian metric associated to his method, in order to study the respective continuous trajectory. That property, which, essentially, means independence under change of coordinates, is also useful to the construction of appropriate metrics underlying interior point methods for LO. We generalize the classical method to obtain invariant Riemannian metric, and we call this generalization *invariant Riemannian metric through H-translation.*

Let S be a differential manifold, $x \in S$ and $H : S \to S$ a differential mapping from S to itself. To construct a invariant Riemannian metric through H-translation the following steps must be followed:

Construction of invariant Riemannian metric through H-translation

1. Define on S a Lie group structure.

2. Given $x \in S$, consider the element $L_{H(x)^{-1}} x$, define a inner product

$$\langle , \rangle_{L_{H(x)^{-1}} x} : T_{L_{H(x)^{-1}} x} S \times T_{L_{H(x)^{-1}} x} S \to I\!R.$$

3. $\forall u, v \in T_x(S)$ define the metric

$$< u, v >_x = \langle d(L_{H(x)^{-1}})_x u, d(L_{H(x)^{-1}})_x v\rangle_{L_{H(x)^{-1}} x}$$

Clearly the metric defined by this method is left invariant at the point x (see 2.1). In the following, the algorithm above will be applied to generate diagonal Riemannian metrics associated to some natural feasible sets for linear optimization problems.

Example 1 *Diagonal metrics on the positive octant $I\!R^n_{++}$.*

Consider the differential manifold $I\!R^n_{++}$, and let $H : I\!R^n_{++} \to I\!R^n_{++}$ be a function such that $H(x) = (h_1(x_1), h_2(x_2), ..., h_n(x_n))$, where $h_i : I\!R_{++} \to I\!R_{++}$ are differentiable functions.
Step 1: defining a Lie group structure.

Let $x = (x_1, x_2, ..., x_n),\ y = (y_1, y_2, ..., y_n) \in R^n_{++}$, set:

$$x \bullet y = (x_1.y_1, x_2.y_2, ..., x_n.y_n)$$

That operation defines a group structure with identity element $e = (1, 1, 1, ..., 1)$ and the inverse of an element $x = (x_1, x_2, ..., x_n)$ given by $x^{-1} = (\frac{1}{x_1}, \frac{1}{x_2}, ..., \frac{1}{x_n})$.
Step 2: defining an inner product at $L_{H(x)^{-1}}x$.

Let $x = (x_1, x_2, ..., x_n) \in I\!R^n_{++}$, then $L_{H(x)^{-1}}x = H(x)^{-1} \bullet x = \left(\frac{x_1}{h_1(x)}, \frac{x_2}{h_2(x)}, ..., \frac{x_n}{h_n(x)}\right)$. As $T_{H(x)^{-1}\bullet x} I\!R^n_{++} = I\!R^n$, we can define a inner product at this point as a Euclidean inner product:

$$\langle v, w\rangle_{L_{H(x)^{-1}}x} = (v, w) = v^T w$$

Step 3: defining the metric for all $x \in I\!R^n_{++}$

First, we shall obtain $d(L_{H(x)^{-1}})$. The application $L_{H(x)} : I\!R^n_{++} \to I\!R^n_{++}$ is by definition: $L_{H(x)}y = H(x) \bullet y = (h_1(x_1)y_1, h_2(x_2)y_2, ..., h_n(x_n)y_n)$. Then $d(L_{H(x)})_y : I\!R^n \to I\!R^n$ can be expressed by the matrix

$$d(L_{H(x)})_y = \mathtt{diag}(h_1(x_1), h_2(x_2), ..., h_n(x_n))$$

and, consequently

$$d(L_{H^{-1}(x)})_y = \mathtt{diag}(\frac{1}{h_1(x_1)}, \frac{1}{h_2(x_2)}, ..., \frac{1}{h_n(x_n)}).$$

Now, $\forall u, v \in T_x(I\!R^n_{++}) = I\!R^n$ we define:

$< u, v >_x = (d(L_{H^{-1}(x)})_x u, d(L_{H^{-1}(x)})_x v) = \sum_{i=1}^{n} \frac{u_i.v_i}{h_i(x_i)^2} = u^T G(x)v$, where

$$G(x) = \mathtt{diag}\left(\frac{1}{h_1^2(x_1)}, \frac{1}{h_2^2(x_2)}, ..., \frac{1}{h_n^2(x_n)}\right).$$

Thus we obtain

$$g_{ij}(x) =< e_i, e_j >_x = \frac{\delta_{ij}}{h_i(x_i)h_j(x_j)}.$$

In particular:

- If $h_i : I\!R_{++} \to I\!R_{++}, h_i(x_i) = x_i^{\frac{r}{2}}$, where r is a non zero scalar, then

$$G(x) = \mathtt{diag}\left(\frac{1}{x_1^r}, \frac{1}{x_2^r}, ..., \frac{1}{x_n^r}\right) = X^{-r}$$

  where $X = \mathtt{diag}(x_1, x_2, ..., x_n)$ This metric, for $r \geq 1$, has been utilized to obtain certain classes of methods in Continuous Optimization, see for example, [4], [13], [17], [18],[19].

- If $h_i : I\!R_{++} \to I\!R_{++}, h_i(x_i) = s_i^{\frac{-r}{2}} x_i^{\frac{r}{2}}$, where $s_i \in I\!R_{++}$ is a fixed point, and r is a non null scalar, then

$$G(x) = \mathtt{diag}\left(\frac{s_1^r}{x_1^r}, \frac{s_2^r}{x_2^r}, ..., \frac{s_n^r}{x_n^r}\right) = S^r X^{-r}$$

  That metric seems to be new in the context of continuous optimization.

Example 2 *Diagonal Riemannian metric on the hypercube.*

Consider the differential sub manifold $C_0^n = \langle 0,1\rangle^n$ of $I\!R^n_{++}$, and let $H : C_0^n \to I\!R^n_{++}$ such that $H(x) = (h_1(x_1), h_2(x_2), ..., h_n(x_n))$, where the $h_i : \langle 0,1\rangle \to I\!R_{++}$ are differentiable functions. Now, we can define the metric on C_0^n as the induced metric of $I\!R^n_{++}$, introduced in the previous example. Thus $\forall u, v \in T_x(C_0^n) = I\!R^n$:

$$< u, v >_x = \sum_{i=1}^{n} \frac{u_i.v_i}{h_i(x_i)^2} = u^T G(x) v.$$

where

$$G(x) = \texttt{diag}\left(\frac{1}{h_1^2(x_1)}, \frac{1}{h_2^2(x_2)}, ..., \frac{1}{h_n^2(x_n)}\right)$$

In particular:

- If $h_i(x_i) = (x_i^r(1-x_i)^r/(x_i^r + (1-x_i)^r))^{1/2}$, for $x_i \in \langle 0,1\rangle,\ i = 1, ..., n$ and r non null, then we have the metric
$$G(x) = X^{-r} + (I - X)^{-r}$$
- If $h_i(x_i) = \sin^2(\pi x_i)/\pi$, for $x_i \in \langle 0,1\rangle$, we have a trigonometric metric
$$\pi^2 \csc^4(\pi x) = \texttt{diag}(\pi^2 \csc^4(\pi x_1), \pi^2 \csc^4(\pi x_2), ..., \pi^2 \csc^4(\pi x_n))$$
- If $h_i(x_i) = x_i^{r/2}(1-x_i)^{r/2}$, for $x_i \in \langle 0,1\rangle$, r non null, we have the metric:
$$X^{-r}(I - X)^{-r}$$

Example 3 *Diagonal Riemannian metric on the product of positive octants .*

Let S_1^n and S_2^m be two Riemannian manifolds with respective inner products $\langle , \rangle_1$ and $\langle , \rangle_2$, and let $S_1 \times S_2$ be the $n + m$ dimension product manifold. Also, we need the projections $\pi^1 : S_1 \times S_2 \to S_1$ and $\pi^2 : S_1 \times S_2 \to S_2$, defined by $\pi^1(p, q) = p$ and $\pi^2(p, q) = q$, where $(p, q) \in S_1 \times S_2$. It is well-known that the dimension of the tangent space at $(p, q) \in S_1 \times S_2$, $T_{(p,q)}(S_1 \times S_2)$, is $n+m$. Denote $P = d\pi^1_{(p,q)}$ and $Q = d\pi^2_{(p,q)}$ the differentials of those projections, then we can define a metric on $S_1 \times S_2$, by introducing the inner product of $u, v \in T_{(p,q)}(S_1 \times S_2)$ as:

$$\langle u, v\rangle_{(p,q)} = \langle P(u), P(v)\rangle_p + \langle Q(u), Q(v)\rangle_q \tag{3.5}$$

Now, consider the Riemannian manifolds $(I\!R^n_{++}, G(x))$, $(I\!R^n_{++}, G(s))$, invariant through H-translation and T-translation respectively, where $H(x) = (h_1(x_1), h_2(x_2), ..., h_n(x_n))$ and $T(s) = (T_1(s_1), T_2(s_2), ..., T_n(s_n))$ for functions $T_i, H_i : I\!R_{++} \to I\!R_{++}$, with metrics expressed by

$$G(x) = \texttt{diag}\left(\frac{1}{h_1(x_1)^2}, \frac{1}{h_2(x_2)^2}, ..., \frac{1}{h_n(x_n)^2}\right)$$

and

$$G(s) = \texttt{diag}\left(\frac{1}{T_1(s_1)^2}, \frac{1}{T_2(s_2)^2}, ..., \frac{1}{T_n(s_n)^2}\right)$$

Using (3.5) we introduce a metric on $I\!R^n_{++} \times I\!R^n_{++}$.
Consider the projections

$$\pi^1 : I\!R^n_{++} \times I\!R^n_{++} \to I\!R^n_{++}, \pi^1(x, s) = x$$

$$\pi^2 : \mathbb{R}^n_{++} \times \mathbb{R}^n_{++} \to \mathbb{R}^n_{++}, \pi^2(x,s) = s.$$

Due to the fact that $T_{(x,s)}(\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}) = T_x\mathbb{R}^n_{++} \times T_s\mathbb{R}^n_{++} = \mathbb{R}^n \times \mathbb{R}^n$, the differentials of the projections at $(x,s) \in \mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$ are respectively:

$$d\pi^1_{(x,s)} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,$$

$$d\pi^1_{(x,s)}(u_x, u_s) = [I_{n\times n} \; 0_{n\times n}] \begin{bmatrix} u_x \\ u_s \end{bmatrix} = u_x$$

$$d\pi^2_{(x,s)} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n,$$

$$d\pi^2_{(x,s)}(u_x, u_s) = [0_{n\times n} \; I_{n\times n}] \begin{bmatrix} u_x \\ u_s \end{bmatrix} = u_s$$

where $u = (u_x, u_s)$, with $u_x \in \mathbb{R}^n$ and $u_s \in \mathbb{R}^n$. Now, we can define for $u = (u_x, u_s)$, $v = (v_x, v_s) \in T_{(x,s)}(\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}) = \mathbb{R}^{2n}$

$$\langle u, v\rangle_{(x,s)} = \langle u_x, v_x\rangle_x + \langle u_s, v_s\rangle_s$$

or

$$\langle u, v\rangle_{(x,s)} = u^T G_1(x) v + v^T G_2(s) v$$

Thus, this inner product defines a Riemannian metric for the "primal-dual" manifold $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$. Taking the canonic basis $\{\partial/\partial x_i\}_{i=1,\dots,n}$ on $\mathbb{R}^n \times \mathbb{R}^n$ we have that the matrix that represents the metric on the product manifold $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$ is:

$$G(x,s) = \begin{bmatrix} G_1(x)_{n\times n} & 0_{n\times n} \\ & \\ 0_{n\times n} & G_2(s)_{n\times n} \end{bmatrix}_{(2n\times 2n)}$$

In particular, if $G_1(x) = X^{-r}$ and $G_2(s) = S^{-r}$ then we have

$$G(x,s) = \begin{bmatrix} X^{-r}_{n\times n} & 0_{n\times n} \\ & \\ 0_{n\times n} & S^{-r}_{n\times n} \end{bmatrix}_{(2n\times 2n)}$$

Remark:

The manner we introduced the construction of invariant Riemannian metrics through H-translation, includes a great number of possibilities to define metrics. In particular, we refer to the class obtained by the Hessian of barrier functions defined by $h_i(x_i) = (1/p_i''(x_i))^{1/2}$ where $p_i : \mathbb{R}_{++} \to \mathbb{R}$ are such that $p_i''(x_i) > 0$. The motivation comes from barrier functions with the form

$$p(x) = \sum_{i=1}^{n} p_i(x_i),$$

see, e. g., (see [14]. pp.15)

4 The Diagonal Riemannian Metric: Geometric Properties

In this section we derive some geometric properties of the general diagonal Riemannian metric

$$G(x) = \texttt{diag}\left(\frac{1}{h_1(x_1)^2}, \frac{1}{h_2(x_2)^2}, \ldots, \frac{1}{h_n(x_n)^2}\right)$$

defined on the positive octant $I\!R^n_{++}$. In this way, we obtain simple expressions to the Christoffel symbols, the equation to obtain explicit geodesic curve is simplified to solve an integral of a certain function. We prove the null curvature of the manifold, present sufficient conditions that assure its completeness, we give explicit expressions for the gradient and Hessian for C^2 functions defined on the manifold. Due to the fact that the metric on the hypercube C^n_0 is induced by the metric on $I\!R^n_{++}$ the results are extendable to the hypercube. Finally, using properties of Riemannian Geometry, ours results are also extended to $I\!R^n_{++} \times I\!R^n_{++}$. To finish the section, we give some examples of explicit geodesics through the introduction of some particular metrics.

1. Christoffel's Symbols.
   We use the relation between metrics and Christoffel's symbols, given in (2.3). If $k \neq m$ then $g^{mk} = 0$, and the expression is reduced to:

$$\Gamma^m_{ij} = \frac{1}{2}\left\{\frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_j} g_{mi} - \frac{\partial}{\partial x_m} g_{ij}\right\} g^{mm}$$

   We consider two cases.
   First case: $i = j$

$$\Gamma^m_{ii} = \frac{1}{2}\left\{\frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_i} g_{mi} - \frac{\partial}{\partial x_m} g_{ii}\right\} g^{mm}$$

   If $m = i$ then

$$\Gamma^m_{ii} = -\frac{1}{h_i(x_i)} \frac{\partial h_i(x_i)}{\partial x_i}$$

   otherwise,

$$\Gamma^m_{ii} = 0.$$

   Second case: $i \neq j$

$$\Gamma^m_{ij} = \frac{1}{2}\left\{\frac{\partial}{\partial x_i} g_{im} + \frac{\partial}{\partial x_j} g_{mi}\right\} g^{mm}$$

   If $m = i$ then $m \neq j$ and:

$$\Gamma^i_{ij} = 0$$

   If $m = j$ then $m \neq i$ and:

$$\Gamma^j_{ij} = 0$$

   If $m \neq i$ and $m \neq j$ then:

$$\Gamma^m_{ij} = 0$$

   In both cases we have

$$\Gamma^m_{ij} = -\frac{1}{h_i(x_i)} \frac{\partial h_i(x_i)}{\partial x_i} \delta_{im}\delta_{ij} \tag{4.6}$$

2. Covariant Derivative and Parallel Transport

Given a vector field $V = \sum_{i=1}^{n} v_i X_i$ on the curve $x(t) = (x_1(t), x_2(t), x_n(t)) \in \mathbb{R}^n_{++}$, where $\{X_i\}_{i=1,\dots,n}$ is a basis for $T_p(\mathbb{R}^n_{++}) = \mathbb{R}^n$. It is well known that

$$\frac{DV}{dt} = \sum_{i=1}^{n} \left(\frac{dv_i}{dt} + \sum_{i,j=1}^{n} v_i \frac{dx_i}{dt} \Gamma_{ij}^k \right) X_k \tag{4.7}$$

Substituting the Christoffel symbols (4.6) in (4.7) we have:

$$\frac{DV}{dt} = \sum_{i=1}^{n} \left(\frac{dv_i}{dt} - \frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} v_i \frac{dx_i}{dt} \right) X_i. \tag{4.8}$$

Now, for $V_0 = (V_1^0, V_2^0, ..., V_n^0) \in T_p(\mathbb{R}^n_{++})$, the differential equation that gets the parallel transport $V(t) = (V_1(t), ..., V_n(t))$ along the curve $x(t)$ is:

$$\frac{dV_i}{dt} - \frac{1}{h_i(x_i)} \frac{\partial(h_i(x_i))}{\partial x_i} V_i \frac{dx_i}{dt} = 0, \quad \forall i = 1, ...n$$

with the condition

$$V(0) = V_0.$$

It is easy to check that this equation is solved by:

$$V(t) = v_i^0 h_i(x_i(t))/h_i(p_i), \forall i = 1, 2, ..., n \tag{4.9}$$

Therefore we can define, in closed form, the parallel transport along the curve $x(t)$ as the application $P_{x(t)} : T_p(\mathbb{R}^n_{++}) \to T_{x(t)}\mathbb{R}^n_{++}$ such that

$$P_{x(t)}(V) = (v_1 h_1(x_1(t))/h_1(p_1), v_2 h_2(x_2(t))/h_2(p_2), ..., v_n h_n(x_n(t))/h_n(p_n))$$

In particular: If $h_i(x_i) = x_i^{r/2}$ ($G(x) = X^{-r}, r \in \mathbb{R}$ non null) then

$$P_{x(t)}(V)_i = v_i x_i^{r/2} / p_i^{r/2} \tag{4.10}$$

If $h_i(x_i) = x_i^{r/2}(1 - x_i)^{r/2}$ ($G(x) = X^{-r}(I - X)^{-r}, r \in \mathbb{R}$ non null) with $x_i \in \langle 0, 1 \rangle$ then

$$P_{x(t)}(V)_i = v_i x_i^{r/2}(1 - x_i)^{r/2} / p_i^{r/2}(1 - p_i)^{r/2} \tag{4.11}$$

If $h_i(x_i) = \sin^2(\pi x_i)/\pi$, for $x_i \in \langle 0, 1 \rangle$, we have

$$P_{x(t)}(V)_i = v_i sin^2(\pi x_i(t))/sin^2(\pi p_i) \tag{4.12}$$

3. Geodesic Equation.

Let $p = (p_1, p_2, ..., p_n) \in \mathbb{R}^n_{++}$ and $v = (v_1, v_2, ..., v_n) \in T_p(\mathbb{R}^n_{++}) = \mathbb{R}^n$ with

$$x : I \to \mathbb{R}^n_{++}; \; x(t) = (x_1(t), x_2(t), ..., x_n(t))$$

such that $x(0) = p$ and $dx(0)/dt = v$.

Substituting the Christoffel's symbols (4.6) in the equation(2.2) we have:

$$\frac{d^2x_i}{dt^2} - \frac{1}{h_i(x_i)}\frac{\partial h_i(x_i)}{\partial x_i}\left(\frac{dx_i}{dt}\right)^2 = 0, i = 1, ..., n \tag{4.13}$$

with initial conditions:

$$\begin{array}{ll} x_i(0) = p_i, & i = 1, ..., n \\ x_i'(0) = v_i, & i = 1, ..., n \end{array}$$

It is easy to check that the differential equation (4.13) is equivalent to:

$$\int \frac{1}{h_i(x_i)} dx_i = a_i t + b_i, \qquad i = 1, 2, ..., n$$

for some constants a_i,and b_i in $I\!R$.

Thus, the unique geodesic of $I\!R^n_{++}$ with metric $G(x)$ is a curve $x(t)$ that solve:

$$\int \frac{1}{h_i(x_i)} dx_i = a_i t + b_i \qquad i = 1, ..., n \tag{4.14}$$

where a_i and b_i are real constants such that:

$$\begin{array}{ll} x_i(0) = p_i, & i = 1, ..., n \\ x_i'(0) = v_i, & i = 1, ..., n \end{array}$$

4. Null Curvature

Given $p \in I\!R^n_{++}$, let $\{X_i\}_{i=1,...,n}$ be a basis for $T_p(I\!R^n_{++}) = I\!R^n$. Then, we can write $X = \sum\limits_{i=1}^{n} u^i X_i$, $Y = \sum\limits_{j=1}^{n} v^j X_j$, $Z = \sum\limits_{k=1}^{n} w^k X_k$. As the curvature tensor R is tri-linear we have:

$$R(X, Y)Z = \sum_{i,j,k} u^i v^j w^k R(X_i, X_j)X_k.$$

From the definition of R (see Section 2) we obtain:

$$R(X_i, X_j)X_k = \nabla_{X_j}(\nabla_{X_i}X_k) - \nabla_{X_i}(\nabla_{X_j}X_k) + \nabla_{[X_i,X_j]}X_k$$

Now, for the Riemannian connection, we have $[X_i, X_j] = 0$, and

$$R(X_i, X_j)X_k = \nabla_{X_j}(\nabla_{X_i}X_k) - \nabla_{X_i}(\nabla_{X_j}X_k)$$

If $i = j$ then $R(X_i, X_j)X_k = 0$. For $i \neq j$, recall, from section 2, that

$$\nabla_{X_i}X_k = \sum_{j=1}^{n} \Gamma^j_{ik} X_j$$

Substituting the Christofell's symbols we get

$$\nabla_{X_i}X_k = \sum_{j=1}^{n}\left(-\frac{1}{h_i(x_i)}\frac{\partial h_i(x_i)}{\partial x_i}\delta_{ij}\delta_{ik}\right)X_j = -\frac{1}{h_i(x_i)}\frac{\partial h_i(x_i)}{\partial x_i}\delta_{ik}X_i \tag{4.15}$$

Therefore:

$$\nabla_{X_j}(\nabla_{X_i}X_k) = \nabla_{X_j}\left(-\frac{1}{h_i(x_i)}\frac{\partial h_i(x_i)}{\partial x_i}\delta_{ik}X_i\right).$$

We have the following cases to analyze. If $i \neq k$, it is immediate that $\nabla_{X_j}(\nabla_{X_i} X_k) = 0, i = 1, ..., n$. Otherwise, if $i = k$, and $j \neq k$, we get, applying (4.15)

$$\nabla_{X_j}(\nabla_{X_k} X_k) = \nabla_{X_j}\left(-\frac{1}{h_k(x_k)}\frac{\partial h_k(x_k)}{\partial x_k}X_k\right) = \frac{1}{h_k(x_k)}\frac{\partial h_k(x_k)}{\partial x_k}\delta_{jk}\frac{1}{h_j(x_j)}\frac{\partial h_j(x_j)}{\partial x_j}X_j = 0$$

Thus

$$\nabla_{X_j}(\nabla_{X_i} X_k) = 0$$

In the same way we have

$$\nabla_{X_i}(\nabla_{X_j} X_k) = 0$$

Those results lead to:

$$R(X_i, X_j)X_k = 0, \quad i, j, k, l = 1, 2, ...n.$$

so $R(X, Y)Z = 0$. Then, the Riemannian manifold $I\!R^n_{++}$, endowed with the metric $G(x)$ has null curvature.

5. Sufficient Conditions for Completeness

   Essential properties as the convergence of algorithms, needs, from the point of view of Riemannian geometry applications, the completeness of the manifold. Oliveira and Cruz Neto [14] have proved a sufficient condition for $I\!R^n_{++}$, endowed with a diagonal metric, which is the Hessian of a certain barrier function. That proof can be generalized to the diagonal metric $G(x)$.

   Theorem 4.1 *Let the manifold $I\!R^n_{++}$ endowed with the invariant diagonal metric $G(x)$. The following is true, for each $i = 1, ..., n$:*

   (a) If $x_i^\alpha \geq \beta h_i(x_i)^2$, with $\alpha \geq 2$ for $x_i \in \langle 0, 1]$, or $\alpha \leq 2$ for $x_i \in \langle 1, \infty\rangle$, for some $\beta > 0$ then, $I\!R^n_{++}$ is isometric with $I\!R^n$. In particular $(I\!R^n_{++}, G(x))$ is complete.

   (b) If $x_i^\alpha \leq \beta h_i(x_i)^2$, with $\alpha < 2$ for $x_i \in \langle 0, 1]$, for some $\beta > 0$, or $x_i^\alpha \geq \beta h_i(x_i)^2$, with $\alpha > 2$ for $x_i \in \langle 0, 1]$, for some $\beta > 0$, then $(I\!R^n_{++}, G(x))$ is incomplete.

6. The Product Manifold $I\!R^n_{++} \times I\!R^n_{++}$.

   We know that the Riemannian manifolds $(I\!R^n_{++}, G_1(x))$, $(I\!R^n_{++}, G_2(s))$ have null curvature. Then, due to a result of Riemannian geometry, the product manifold $\left(I\!R^n_{++} \times I\!R^n_{++}, G(x, s)\right)$ has null curvature. Moreover, if the manifolds $I\!R^n_{++}$ and $I\!R^n_{++}$ are complete with geodesic curves γ_1 e γ_2, then the product manifold $I\!R^n_{++} \times I\!R^n_{++}$ is complete with geodesic curves $\gamma_1 \times \gamma_2$.

7. Gradient and Hessian

   Let the Riemannian manifold $I\!R^n_{++}$ endowed with the metric $G(x) = \texttt{diag}(1/h_1(x_1)^2, ..., 1/h_n(x_n)^2)$, where $h_i : I\!R_{++} \to I\!R_{++}$, $i = 1, 2, ..., n$, are differentiable functions, and $f : I\!R^n_{++} \to I\!R$

   (a) The gradient and the Hessian of f in the Riemannian manifold $I\!R^n_{++}$, denoted, respectively, by $\nabla_{I\!R^n_{++}} f$ and H^f_x are:

$$\nabla_{I\!R^n_{++}} f(x) = G(x)^{-1} f'(q) = \left(h_1^2(x_1)\frac{\partial f}{\partial x_1}(x), ..., h_n^2(x_n)\frac{\partial f}{\partial x_n}(x)\right) \qquad (4.16)$$

$$H_x^f = f''(x) + G^{\frac{1}{2}}(x)(G^{-\frac{1}{2}}(x))'\mathcal{F}'(x) \tag{4.17}$$

where:
$\mathcal{F}'(x) = \texttt{diag}\left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, ..., \frac{\partial f(x)}{\partial x_n}\right)$ and $f''(x) = \texttt{diag}\left(\frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_2^2}, ..., \frac{\partial^2 f}{\partial x_n^2}\right)$.

In particular:

i. If $h_i(x_i) = x_i^{\frac{r}{2}}, r \in I\!R$ then

$$\nabla_{I\!R^n_{++}} f(x) = \sum_{i=1}^{n} x_i^r \frac{\partial f}{\partial x_i}(x) = X^r f'(x)$$

$$H_x^f = f''(x) + \frac{r}{2} X^{-1}\mathcal{F}'(x)$$

(b) Similarly, the gradient and the Hessian of f in the manifold $C_0^n = \langle 0, 1\rangle^n$ have the same expressions above, given in (4.16) and (4.17).
In particular:

i. If the metric is given by $G(x) = X^{-r} + (I - X)^{-r}$ then

$$\nabla_{C_0^n} f(x) = (X^{-r} + (I - X)^{-r})^{-1} f'(x)$$

$$H_x^f = f''(x) + \frac{r}{2}[X^{-r} + (I - X)^{-r}]^{-1}[X^{-r-1} - (I - X)^{-r-1}]\mathcal{F}'(x)$$

ii. If $G(x) = \pi^2 \csc^4(\pi x)$ then

$$\nabla_{C_0^n} f(x) = \frac{1}{\pi^2} \sin^4(\pi x) f'(x)$$

$$H_x^f = f''(x) + 2\pi^3 X \cot(\pi x)\mathcal{F}'(x).$$

where $\cot(x) = \texttt{diag}(\cot(x_1), ..., \cot(x_n))$

iii. If the metric is $G(x) = X^{-r}(I - X)^{-r}$ then

$$\nabla_{C_0^n} f(x) = X^r (I - X)^r f'(x).$$

$$H_x^f = f''(x) + \frac{r}{2}[X^{-1} - (I - X)^{-1}]\mathcal{F}'(x)$$

(c) Let $I\!R^n_{++} \times I\!R^n_{++}$ be the product manifold endowed with the Riemannian metric under consideration, then we get

$$\nabla_{I\!R^n_{++} \times I\!R^n_{++}} f(x, s) = \begin{bmatrix} G_1(x)_{n\times n} & O \\ 0 & G_2(s)_{n\times n} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial s} \end{bmatrix}$$

(d) If we consider the sub manifold $M = \{x \in I\!R^n_{++},\ Ax = b\}$ be endowed with the metric $G(x)$ induced from $I\!R^n_{++}$ then

$$\nabla_M f(x) = P_M(x)\nabla_{I\!R^n_{++}} f(x)$$

where $P_M(x) = (I - G(x)^{-1}A^T(AG^{-1}(x)A^T)^{-1}A)$, the known $G(x)$ projection operator.

(e) Consider the sub manifold $M = \{x \in I\!R^n,\ l(x) = 0\}$, with induced metric $G(x)$ from $I\!R^n$, then

$$\nabla_M f(x) = P_M(x)\nabla_{I\!R^n} f(x)$$

where $P_M(x) = \Big(I - G(x)^{-1}J_l(x)^T(J_l(x)G^{-1}(x)J_l(x)^T)^{-1}J_l(x)\Big)$.

4.1 Examples of Explicit Geodesic Curves:

1. Consider the manifold $(\mathbb{R}^n_{++}, X^{-r})$. The geodesic curves $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ defined in this manifold such that $x(0) = p = (p_1, p_2, ..., p_n)$ and $x'(0) = v = (v_1, v_2, ..., v_n)$ are:

   (a) for $r \neq 2$,

   $$x_i(t) = \left(\frac{2-r}{2}\right)^{\frac{2}{2-r}} \left(\frac{v_i}{p_i^{\frac{r}{2}}}t + \frac{2}{2-r}p_i^{1-\frac{r}{2}}\right)^{\frac{2}{2-r}}, \ i = 1, 2..., n.$$

   (b) for $r = 2$,

   $$x_i(t) = p_i exp\left(\frac{v_i}{p_i}t\right), \ i = 1, 2..., n.$$

   We observe that when $r = 2$, the geodesic curve $x(t)$ is defined for all $t \in \mathbb{R}$. Moreover, the Riemannian distance from $p = x(0)$ to $q = x(t_0)$, $t_0 > 0$, is given by:

   $$d(p, q) = \int_0^{t_0} \|x'(t)\| dt = \left\{\sum_{i=1}^n \left[\ln\left(\frac{q_i}{p_i}\right)\right]^2\right\}^{\frac{1}{2}}$$

   Therefore, $(\mathbb{R}^n_{++}, X^{-2})$ is complete with null curvature.

2. Consider the manifold $(C^n_0, \pi^2 \csc^4(\pi x))$, where $h_i(x_i) = \frac{1}{\pi}\sin^2(\pi x_i)$. The geodesic curve $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ defined in this manifold, such that $x(0) = p = (p_1, p_2, ..., p_n)$ and $x'(0) = v = (v_1, v_2, ..., v_n)$ is:

   $$x_i(t) = \frac{1}{\pi}\arctan\Big(-\pi \csc^2(\pi p_i)v_i t + \cot(\pi p_i)\Big), \ i = 1, 2..., n. \tag{4.18}$$

   The geodesic curve is well defined for all $t \in \mathbb{R}$. The Riemannian distance from $p = x(0)$ to $q = x(t_0)$, $t_0 > 0$, is given by:

   $$d(p, q) = \int_0^{t_0} \|x'(t)\| dt = \left\{\sum_{i=1}^n [\cot(\pi q_i) - \cot(\pi p_i)]^2\right\}^{\frac{1}{2}} \tag{4.19}$$

   Thus, this manifold is complete with null curvature.

3. Finally, consider $(C^n_0, X^{-2}(I - X)^{-2})$ with $h_i(x_i) = x_i(1 - x_i)$. The geodesic curve $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ defined in this manifold, such that $x(0) = p = (p_1, p_2, ..., p_n)$ and $x'(0) = v = (v_1, v_2, ..., v_n)$ is:

   $$x_i(t) = \frac{1}{2}\left\{1 + \tanh\left(\frac{1}{2}\frac{v_i}{p_i(1-p_i)}t + \frac{1}{2}\ln\frac{p_i}{1-p_i}\right)\right\} \ i = 1, 2..., n. \tag{4.20}$$

   where $\tanh(z) = (e^z - e^{-z})/(e^z + e^{-z})$ is the hyperbolic tangent function. The geodesic curve is well defined for all $t \in \mathbb{R}$. The Riemannian distance from $p = x(0)$ to $q = x(t_0)$, $t_0 > 0$, is given by:

   $$d(p, q) = \int_0^{t_0} \|x'(t)\| dt = \left\{\sum_{i=1}^n \left[\ln\left(\frac{q_i}{1-q_i}\right) - \ln\left(\frac{p_i}{1-p_i}\right)\right]^2\right\}^{\frac{1}{2}} \tag{4.21}$$

   In the same way, this manifold is complete with null curvature.

5 Applications

In this section we present some applications of Riemannian geometry to solve optimization problems. The first class we are interested is min $f(x)$ s.t: $0 \leq x_i \leq 1,\ i = 1, 2, ..., n$, where f can be a differentiable or non-differentiable function. We show some new examples of (explicit) descent geodesic algorithms. The convergence results for those algorithms need convexity hypothesis, and have been proved for general Riemannian metrics by da Cruz Neto et al. [2], for the differentiable case, and Ferreira and Oliveira [5], for non-smooth functions. Besides, we observe, that, for the smooth case, the application of Luenberger's idea, see [9], in his descent geodesic gradient projection method, the ensuring of linear rate of convergence for that method, allows, under some reasonable hypothesis, to guarantee the same property for the usual line search corresponding algorithm. A second application is the using of the proximal algorithm for the same problem above, where f can be a non-differentiable function. The convergence theory is given in [6]. To finish the section, we introduce two barrier functions to solve the problem, min $f(x)$ s.to $Ax = b$, $0 \leq x_i \leq 1,\ i = 1, 2, ..., n$. We prove that those barriers have self-concordant properties, one of them being a $(9/4)n$-self-concordant barrier, so we have a new class of polynomial algorithms, with identical complexity to classic logarithmic barrier.

5.1 Application 1: Geodesic Algorithms

Some examples of geodesic methods to solve minimization problems on the positive octant and the unitary simplex have been reported by Cruz Neto and Oliveira [1]. In this subsection we get more examples, for the minimization on the hypercube.

Let the problem:

$$\begin{array}{l} \min\ f(x) \\ s.to \\ 0 \leq x_i \leq 1 \quad i = 1, 2, ..., n \end{array} \tag{5.22}$$

where $f : C^n \to I\!R$, $x = (x_1, x_2, ..., x_n) \in \langle 0, 1\rangle^n$.

We take $\langle 0, 1\rangle^n$, and two Riemannian manifolds: the first endowed with the metric $G(x) = X^{-2}(I - X)^{-2}$, and the second, with $M(x) = \pi^2 \csc^4(\pi x)$. Before we present the algorithms we will deduce some geometric properties that are useful to obtain convergence properties.

- the manifolds $\langle 0, 1\rangle^n$ are geodesic convex. Indeed, let $p, q \in \langle 0, 1\rangle^n$, then, for $G(x)$, the unique geodesic $x : [0, 1] \to \langle 0, 1\rangle^n$, joining p and q, where $x(0) = p$ and $x(1) = q$ is:

  $$x_i(t) = \frac{1}{2}\left[1 + \tanh\left(\frac{1}{2}\left\{\left[\ln\left(\frac{q_i}{1 - q_i}\right) - \ln\left(\frac{p_i}{1 - p_i}\right)\right]t + \ln\left(\frac{p_i}{1 - p_i}\right)\right\}\right)\right],$$

  and, for $M(x)$,

  $$x_i(t) = \frac{1}{\pi}\texttt{arccotan}[\cot\left[(\cot(\pi q_i) - \cot(\pi p_i))t + \cot(\pi p_i)\right],$$

  It is easy to check, for both cases, that $0 < x_i(t) < 1, i = 1, 2, ..., n$. Thus, $\langle 0, 1\rangle^n$ is geodesic convex.

- The manifolds are connected (trivial)

- The manifolds are complete and have null curvature, with explicit geodesics (see 4.18 e 4.20)

- The gradient and the Hessian of differentiable functions defined on the manifold are explicit (see 4.16 and 4.17).

5.1.1 Differentiable Case: Cauchy algorithm

We present the geodesic descent algorithm, that works essentially as follows. In the first step, it takes the direction of the negative gradient of the function f, given by $d = -\nabla_{C_0^n} f(x)$ (the gradient obtained through the specific Riemannian metric). In the second step, it takes the geodesic starting at point p with direction d, and the last step, makes a search along the geodesic. We consider Armijo and fixed step search. We guarantee the weak convergence (that is, the geodesic distance between two consecutive iterations converge to zero) for arbitrary differentiable functions, and global and linear rate of convergence under some convexity conditions for the function f.

Algorithm A: Cauchy algorithm with Armijo search

1. Given a tolerance $\epsilon > 0$, $x^k = (x_1^k, x_2^k, ..., x_n^k)$, $k \geq 0$, feasible, compute the metric $G(x^k) = X_k^{-2}(I - X_k)^{-2}$ ($M(x^k) = \pi^2 \csc^4(\pi x^k)$ respectively), where $X_k = \mathtt{diag}(x_1^k, x_2^k, ..., x_n^k)$

2. Compute the descent direction $d^k = -G(x^k)^{-1} f'(x^k) = -X_k^2 (I - X_k)^2 f'(x^k)$ ($d^k = -M(x^k)^{-1} f'(x^k) = -\frac{1}{\pi^2} \sin^4(\pi x^k) f'(x^k)$ respectively)

3. Let the unique geodesic $x(t)$, $t \geq 0$, such that $x(0) = x^k$ and $x'(0) = d^k$, given by:

$$x_i^k(t) = \frac{1}{2}\left\{1 + \tanh\left(\frac{1}{2}\frac{d_i^k}{x_i^k(1 - x_i^k)}t + \frac{1}{2}\ln\frac{x_i^k}{1 - x_i^k}\right)\right\},\ i = 1, 2..., n.$$

$$\left(x_i^k(t) = \frac{1}{\pi}\arctan\left(-\pi\csc^2(\pi x_i^k) d_i^k t + \cot(\pi x_i^k)\right),\ i = 1, 2..., n \text{ respectively}\right)$$

4. Choose $t_k = 2^{-i_k}\bar{\text{t}}$, where $\bar{\text{t}} > 0$ is given, i_k is the least positive natural number such that

$$f(x(t_k)) \leq f(x^k) - \beta t_k^2 \|d^k\|^2$$

where $\beta \in \langle 0, 1\rangle$

5. Make $x^{k+1} = x^k(t_k)$, and compute the geodesic distance between the points x^k and x^{k+1}, as:

$$d(x^k, x^{k+1}) = \left\{\sum_{i=1}^{n}\left[\ln\left(\frac{x_i^{k+1}}{1 - x_i^{k+1}}\right) - \ln\left(\frac{x_i^k}{1 - x_i^k}\right)\right]^2\right\}^{\frac{1}{2}}$$

$$\left(d(x^k, x^{k+1}) = \left\{\sum_{i=1}^{n}\left[\cot(\pi x_i^{k+1}) - \cot(\pi x_i^k)\right]^2\right\}^{\frac{1}{2}}, \text{respectively}\right)$$

6. Stop test: if $||d(x^k, x^{k+1})|| < \epsilon$, stop. Otherwise, make $x^k \leftarrow x^{k+1}$ and return the step 1.

Convergence Results Algorithm B: Cauchy algorithm with fixed step

This algorithm is analogous to the previous one, except the step 4, which is substituted by: Given $\delta_1 > 0$ and $\delta_2 > 0$ such that $\delta_1 \Gamma + \delta_2 < 1$, choose

$$t_k \in \left(\delta_1, \frac{2}{\Gamma}(1 - \delta_2)\right)$$

where Γ is the Lipschitz constant associated to $\nabla_{C_0^n} f(x)$, that is, Γ satisfies the following property: for any $p, q \in C_0^n$ and any geodesic segment $\alpha : [0, a] \to C_0^n$ joing p and q we have

$$|\nabla_{C_0^n} f(\alpha(t)) - P_\alpha(\nabla_{C_0^n} f(p))| \leq \Gamma d(\alpha(0), \alpha(t)) \tag{5.23}$$

for any $t \in [0, a]$, where P_α is the parallel transport, see 4.11, and 4.12.

We have the following results, all of them, proved for general metrics in Cruz Neto et al [2](We observe that for the Algorithm B the objective function must verify the property (5.23))

Theorem 5.1 *(Theorem 5.1 [2])*
Let $f \in C^1$, and $\{x^k\}$ be a sequence of points generated by Algorithm A, then

1. *There exists a constant β such that*
$$f(x^{k+1}) \leq f(x^k) - \beta t_k^2 \|d^k\|^2 \tag{5.24}$$
In particular $f(x^k)$ is non-increasing.

2. *The sequence $\{x^k\}$is weakly convergent, in the following sense,*
   (a) *$\{x^k\}$ is bounded.*
   (b) $\lim\limits_{k\to\infty} d(x^{k+1}, x^k) = 0$
   (c) *Any accumulation point of $\{x^k\}$ is a critical point of f.*

Theorem 5.2 *(Theorem 5.3 [2])*
Let $f \in C^1$ be a geodesic convex function, then the sequence $\{x^k\}$ converges globally to a minimum point, should it exist.

Theorem 5.3 *(Theorem 5.4 [2])*
Let $f \in C^2$ and strong geodesic convex, that is, $a \leq H^f \leq b$, where H^f is the Hessian of f, and $0 < a < b$. Then the convergence rate of Cauchy algorithm with Armijo search is given by the Kantorovich ratio.

5.1.2 Non differentiable case: sub gradient algorithm

The sub gradient algorithm in Riemannian manifolds is a natural extension of the sub gradient algorithm, introduced by Shor, [20].

Consider that the function f in problem (5.22) is geodesic convex on the Riemannian manifold $\langle 0, 1\rangle^n$ with metric $G(x) = X^{-2}(I - X)^{-2}$ ($M(x) = \pi^2 \csc^4(\pi x)$,respectively). Let O denotes the set of minimizers of f, $f^* = inf\, f(x)$ is its minimum value and $\partial f(x)$ is the sub differential of f at a point $x \in \langle 0, 1\rangle^n$. The problem is to estimate f^* and also to find a point in O, if such point exist, that is, if $O \neq \emptyset$.

Algorithm C
Given a sequence $\{t_k\}$ of real numbers with $t_k > 0$, $k = 1, 2, ...$

1. Initialize. Choose $x_1 \in \langle 0, 1\rangle^n$ and obtain $s_1 \in \partial f(x_1)$. Make $k = 1$

2. If $s_k = 0$, stop. Otherwise, compute the unique geodesic $x(t)$ such that $x(0) = x^k$ and $x'(0) = d^k = -\frac{s_k}{||s_k||}$, as:
$$x_i^k(t) = \frac{1}{2}\left\{1 + \tanh\left(\frac{1}{2}\frac{d_i^k}{x_i^k(1 - x_i^k)}t + \frac{1}{2}\ln\frac{x_i^k}{1 - x_i^k}\right)\right\},\ i = 1, 2..., n.$$
$$\left(x_i^k(t) = \frac{1}{\pi}\arctan\left(-\pi\csc^2(\pi x_i^k)d_i^k t + \cot(\pi x_i^k)\right),\ i = 1, 2..., n,\ \text{respectively}\right)$$

3. Make $x^{k+1} = x_k(t_k)$

4. Compute $s_{k+1} \in \partial f(x^{k+1})$, and return to step 2.

Convergence Result

The following result is an adaptation to our Riemannian manifolds on $\langle 0,1\rangle^n$, with metric $G(x) = X^{-2}(I-X)^{-2}$ ($M(x) = \pi^2 \csc^4(\pi x)$, respectively) from a general theorem obtained by Ferreira and Oliveira [5].

Theorem 5.4 *(Teorema 5.1 [5]). Let f be a convex function on the manifold $\langle 0,1\rangle^n$ and $\{x^k\}$ a sequence of points generated by the algorithm B. If the sequence $\{t_k\}$, with $t_k > 0$, $k = 1, 2...$ is chosen to satisfy:*

$$\sum_{k=0}^{\infty} t_k = \infty;$$

$$\sum_{k=0}^{\infty} t_k^2 < \infty$$

then, inf $f(x^k) = f^*$, *when $k \to \infty$. In addition, if $O \neq \emptyset$, then the sequence $\{x^k\}$ converges to a point $x^* \in O$.*

5.2 Application 2: Proximal Point Algorithm

Consider the problem given by (5.22). To each $x \in \langle 0,1\rangle^n$, the Moreau-Yosida regularization to f, with $\beta > 0$ is given by:

$$\begin{array}{rl} f_\beta(x) = & \min \{f(y) + \frac{\beta}{2} d^2(x,y)\} \\ & s.to \\ & y \in \langle 0,1\rangle^n \end{array} \tag{5.25}$$

In this application, the distance d can be chosen as (4.19), or, (4.21).

Definition 5.1 *The point $\bar{x} = \bar{x}\ (x) = \arg\min f_\beta(x)$, is called the proximal point of x with respect to β, f and d^2*

The proximal point algorithm generates, for a starting point $x^0 \in \langle 0,1\rangle^n$, a sequence $\{x_k\} \subset \langle 0,1\rangle^n$ through the iteration:

$$\begin{array}{rl} x^{k+1} = & \arg\min \{f(y) + \frac{\beta^k}{2} d^2(x^k,y)\} \\ & y \in \langle 0,1\rangle^n \end{array} \tag{5.26}$$

where β^k satisfies $\sum\limits_{k=0}^{\infty} \frac{1}{\beta^k} = \infty$.

Based on the preview results on null curvature, completeness of the manifold $\langle 0,1\rangle^n$, and geodesic distance, we can present the proximal point algorithm.

Algorithm D

Given a starting point $x^0 \in \langle 0,1\rangle^n$ and $\beta^0 > 0$. Choose the sequence β^k such that $\sum\limits_{k=0}^{\infty} 1/\beta^k = \infty$;.

1. Make $k = 0$

2. (Stop criterion) check the optimality of x^k.

3. Compute

$$x^{k+1} = \begin{array}{l} \arg\min \{f(y) + \frac{\beta^k}{2} d^2(x^k, y)\} \\ y \in \langle 0,1\rangle^n \end{array} \tag{5.27}$$

4. Update β^k, k, x^k, and return to step 2.

Convergence result

Theorem 5.5 *(from Ferreira and Oliveira [6]) If f is geodesic convex on the manifold $\langle 0,1\rangle^n$ and there exists an optimal solution then, the sequence generates by algorithm C converges globally to the minimum of the problem.*

An important fact is that this algorithm does not need the geodesic.

5.3 Application 3: Barriers for Central Path Following Methods

In this section, motivated by the study of diagonal metrics, we introduce two new barrier functions, one of them being a self-concordant barrier. The problem we are interested is

$$\begin{array}{l} \min\ f(x) \\ s.to \\ Ax = b \\ 0 \leq x_i \leq 1 \quad i = 1, 2, ..., n \end{array} \tag{5.28}$$

where $f : C^n \to I\!R$ is a differentiable function, $A \in I\!R^{m\times n}$ is an $m \times n$, $m < n$, matrix with full range.

5.3.1 Self-concordant Properties

Consider the unitary hypercube $C^n = [0,1]^n$ as the environment space. Let C_0^n be the interior of C^n. In C_0^n we introduce two barrier functions:

$$b_1(x) = -\frac{1}{6}\sum_{i=1}^{n}[4\ln(\sin\pi x_i) - \cot^2(\pi x_i)],$$

$$b_2(x) = \sum_{i=1}^{n}(2x_i - 1)[\ln x_i - \ln(1 - x_i).]$$

The first and the second-order derivatives are, respectively:

$$b_1'(x)_i = -\pi\left(\cot(\pi x_i) + \frac{1}{3}\cot^3(\pi x_i)\right),\ i = 1, ..n.$$

$$b_1''(x) = \pi^2\csc^4(\pi x)$$

$$b_2'(x)_i = 2\left[\ln x_i - \ln(1 - x_i)\right] + (2x_i - 1)\left[\frac{1}{x_i(1 - x_i)}\right],\ i = 1, ..n.$$

$$b_2''(x) = X^{-2}(I - X)^{-2}$$

Clearly, both Hessians are positive definite, so the respective barrier functions are strictly convex. In addition, if $x \to \partial(C^n)$ (x approaches the boundary) then $b_1(x) \to \infty$ and $b_2(x) \to \infty$.

Definition 5.2 *(see[11]) Let E be an open convex finite dimensional vector space, $Q \subset E$ and $B : Q \to I\!R$ a function. B is called self-concordant on Q with parameter $a > 0$ (a-self-concordant) if:*

1. *$B \in C^3$*

2. *B is a convex function on Q.*

3. *For any $x \in Q$ and $h \in E$:*

$$|\nabla^3 B(x)[h,h,h]| \leq 2a^{-\frac{1}{2}}(h^T \nabla^2 B(x) h)^{\frac{3}{2}}$$

If, furthermore, there exists $c > 0$ such that B satisfied:

$$|\nabla B(x)[h]| \leq c^{\frac{1}{2}}(h^T \nabla^2 B(x) h)^{\frac{1}{2}}$$

then, B is called self-concordant barrier with parameter c (c-self-concordant barrier).

Due to the stability of that property wit respect to direct product, see Proposition (2.1.1) in [11], we restrict our analysis to a general term of the barrier:

$$b_1(z) = -\frac{1}{6}\{4\ln(\sin \pi z) - \cot^2(\pi z)\}$$

$$b_2(z) = (2z-1)[\ln z - \ln(1-z)]$$

Theorem 5.6

1. *The function b_1 is a $1/4$ self-concordant function, but it is not self-concordant barrier.*

2. *The function b_2 is a 1-self-concordant function and a $9/4$-self-concordant barrier.*

Proof.
First, we prove 1. Let $z \in \langle 0, 1\rangle$. We have $b_1'(z) = -\{\cot(\pi z) + \cot^3(\pi z)/3\}/\pi$, $b_1''(z) = \csc^4(\pi z)$, and, $b_1'''(z) = -4\pi^3 \csc^4(\pi z)\cot(\pi z)$. Let $a = 1/4$. Then

$$\frac{|b_1'''(z)|}{2a^{\frac{-1}{2}}(b_1''(z))^{\frac{3}{2}}} = \frac{4\pi^3 \csc^4(\pi z)|\cot(\pi z)|}{2(2\pi^3)(\csc^4(\pi z))^{\frac{3}{2}}} = |\sin(\pi z)\cos(\pi z)| \leq 1.$$

This proves that b_1 is 1/4-self-concordant. To prove that it is not a self-concordant barrier, let c be an arbitrary positive real number. Then

$$\frac{|b_1'(z)|}{c\sqrt{b_1"(z)}} = \frac{\pi|\cot(\pi z) + \frac{1}{3}\cot^3(\pi z)|}{c\pi(\csc^4(\pi z))^{\frac{1}{2}}} = \frac{1}{3c}|\sin(\pi z)\cos(\pi z)||3 + \cot^2(\pi z)|$$

Now, it is easy to show that the last term is unbounded at the middle point $z = 1/2$, so, b_1 is not a self-concordant barrier.
Next, we prove 2. Let $z \in \langle 0, 1\rangle$. We have $b_2'(z) = 2[\ln z - \ln(1-z)] + (2z-1)[z(1-z)]^{-1}$, $b_2''(z) = z^{-2}(1-z)^{-2}$ and $b_2'''(z) = 2(2z-1)z^{-3}(1-z)^{-3}$. As a consequence,

$$\frac{|b_2'''(z)|}{2(b_2''(z))^{\frac{3}{2}}} = |2z-1| \leq 1,$$

therefore b_2 is 1-self- concordant. To show that it is a 9/4-self-concordant barrier, take:

$$\frac{|b_2'(z)|}{c^{\frac{1}{2}}(b_2''(z))^{\frac{1}{2}}} = \frac{|2\ln(\frac{z}{1-z}) + (2z-1)\frac{1}{z(1-z)}|}{c^{\frac{1}{2}}\frac{1}{z(1-z)}} = \frac{1}{c^{\frac{1}{2}}}\left|2z(1-z)\ln\left(\frac{z}{1-z}\right) + 2z - 1\right|,$$

which gives:

$$\frac{|b_2'(z)|}{c^{\frac{1}{2}}(b_2''(z))^{\frac{1}{2}}} \leq \frac{1}{c^{\frac{1}{2}}}\left\{\left|2z(1-z)\ln\left(\frac{z}{1-z}\right)\right| + |2z-1|\right\}$$

In the right hand side, the analysis of the critical or extremum points of the first expression, furnishes the following results: for $z = 0$, $z = 1/2$, and $z = 1$, its value is zero. The approximate critical points $z = 0.176041$ and $z = 0.823959$, lead to a value strictly smaller than 2×0.45. As $|2z - 1| \leq 1$, the result follows. ∎

Additionally, due to the stability of that property with respect to the direct product, we have $b_2(x)$ a $(9/4)n$-self-concordant barrier.

6 Conclusions

We proposed a generalization of the method to obtain invariant Riemannian metrics. It has a wide potential to get new algorithms for continuous optimization. Furthermore, using the hypercube as a model in $I\!R^n$, we present some new examples of geodesic methods, we introduce two barrier functions with self-concordant properties . Specially, the barrier $\sum_{i=1}^{n}(2x_i - 1)[\ln x_i - \ln(1-x_i)]$ is a $(9/4)n-$self-concordant barrier. Therefore, using the results of Nesterov and Nemerovski [11] we can obtain polynomial algorithms, to solve the problem of min $c^T x$ s.to $Ax = b$, $0 \leq x_i \leq 1$, $i = 1, 2, ..., n$, with similar complexity as the classic logarithmic barrier. Forthcoming papers exploit those ideas in a context of linear programming, [15], and semidefinite programming, [16].

References

[1] J. X. da Cruz Neto and P. R. Oliveira, Geodesic Methods in Riemannian Manifolds, Technical Report ES-352/95, Systems Engineering and Computer Sciences, PESC/COPPE, Federal University of Rio de Janeiro, 1995.

[2] J. X.da Cruz Neto, L.L. de Lima and P. R. Oliveira, Geodesic Algorithms in Riemannian Geometry, *Balkan Journal of Geometry and its Applications (BJGA)*, v. 3, n.2, 1998, pp. 89-100.

[3] M. P. do Carmo, Riemannian Geometry, Bikhausen, Boston, 1992.

[4] D. den Hertog, Interior Point Approach to Linear, Quadratic and Convex Programming, 1 ed. Kluwer Academic Publishers, 1992.

[5] O. P. Ferreira and P. R. Oliveira, Sub gradient Algorithm on Riemannian Manifolds, Journal of Optimization Theory and Application, v. 97, n.1 (April), 1998, pp. 93-104.

[6] O. P. Ferreira and P. R. Oliveira, Proximal Point Algorithm on Riemannian Manifolds, Optimization, v. 51, n.2, 2002, pp. 257-270. 1955.

[7] N. Karmarkar, A New Polynomial-Time Algorithm for Linear Programming, Combinatorica, v. 4, 1984, pp. 373-395.

[8] N. Karmarkar, Riemannian Geometry Underlying Interior-Point Methods for Linear Programming, Contemporary Mathematics, v. 114, 1990, pp. 51-75.

[9] D. G. Luenberger, The Gradient Projection Method Along Geodesics, *Management Science*, v. 18, n. 11, 1972, pp. 620-631.

[10] J. L. Nazareth, The Newton-Cauchy framework. A unified approach to unconstrained nonlinear minimization, Springer-Verlag, Berlin, Heidelberg, 1994.

[11] Y. E. Nesterov and A. N. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming, Philadelphia, Society for Industrial and Applied Mathematics, 1994.

[12] Y. E. Nesterov and M. J. Todd, On the Riemannian Geometry Defined by Self-Concordant Barrier and Interior-Point Methods, Foundations of Computational Mathematics, v. 2, 2002, pp. 333-361.

[13] G. L. Oliveira and P. r. Oliveira, A New Class of Interior Point Methods for Optimization Under Positivity Constraints. Technical Report, Systems Engineering and Computer Science, PESC/COPPE, Federal University of Rio de Janeiro, 2002.

[14] P. R. Oliveira and J. X. da Cruz Neto, A Unified View of Primal Methods through Riemannian Metrics, Tecnical Report, PESC/COPPE, Federal University of Rio de Janeiro. Engenharia de Sistemas, 1995.

[15] E. A. Papa Quiroz and P. Roberto Oliveira, A new self-concordant barrier for the hypercube, Working paper.

[16] E. A. Papa Quiroz and P. Roberto Oliveira, A new self-concordant barrier for a class of semidefinite programming problems, Working paper.

[17] G. J. Pereira and P. R. Oliveira, A new Class of Proximal Algorithms for the Nonlinear Complementary Problems, Technical Report ES-560/01, PESC/COPPE, Federal University of Rio de Janeiro. Engenharia de Sistemas, 2001.

[18] A. W. M. Pinto, P. R. Oliveira and J. X. da Cruz Neto, A New Class of Potential Affine Algorithms for Linear Convex Programming, Technical Report, Systems Engineering and Computer Science, PESC/COPPE, Federal University of Rio de Janeiro, 2000.

[19] R. Saigal, The primal power affine scaling method, Tech. Report, No 93-21, Dep. Ind. and Oper. Eng., University of Michigan, 1993.

[20] N. Z. Shor, Minimization Methods for Nondifferentiable Functions, Springer Verlag, Berlin, Germany, 1985.

[21] J. S. Wright, Primal-Dual Interior Point Methods, 1 ed. Philadelphia, SIAM, 1997.