A class of self-concordant functions on Riemannian manifolds

Gabriel Bercu and Mihai Postolache

Abstract. The notion of self-concordant function on Euclidean spaces was introduced and studied by Nesterov and Nemirovsky [6]. They have used these functions to design numerical optimization algorithms based on interior-point methods ([7]). In [12], Constantin Udrişte makes an extension of this study to the Riemannian context of optimization methods. In this paper, we use a decomposable function to introduce a new class of self-concordant functions, defined on Riemannian manifolds endowed with metrics of diagonal type. While §1 is introductory in nature, §2 contains our results. We state and prove sufficient conditions for a function to be self-concordant and make two case studies. Examples we found could be used as self-concordant functions to design Newton-type algorithms on smooth manifolds in the sense of Jiang, Moore and Ji [5]. We also solve a very important problem in Riemannian geometry, rised by Professor Constantin Udrişte during the preparation of this paper, regarding the existence of the metric generated by a function which is self self-concordant.

M.S.C. 2000: 53C05.
Key words: self-concordant function, Riemannian manifold, differential inequality.

1 Introduction

Nesterov and Nemirovsky [6] showed that the logarithmic barrier functions for the following problems are self-concordant: linear and convex quadratic programming with convex quadratic constraints, primal geometric programming, matrix norm minimization etc. In [4], D. den Hertog proved that the logarithmic barrier function satisfies the condition to be self-concordant for other important classes of problems.

Many optimization problems can be better stated on manifolds rather than Euclidean space, for example, Newton type methods, [5], or interior-point method in the sense of D. den Hertog [3], [4]. Therefore, it is natural to make a study of self-concordant functions on Riemannian manifolds.

In [12], Constantin Udrişte refers to the general framework of the logarithmic barrier method for smooth convex programming on Riemannian manifolds and shows

Balkan Journal of Geometry and Its Applications, Vol.14, No.2, 2009, pp. 13-20.

that the central path is a minus gradient line and gives the Riemannian generalization for some remarkable results of Nesterov and Nemirovsky. In [5], it is proposed a damped Newton algorithm for optimization of self-concordant functions.

We introduce a class of self-concordant functions defined on the Riemannian manifold $M = \mathbb{R}^n_+$, endowed with the diagonal metric

$$g(x^1, \dots, x^n) = \begin{pmatrix} \frac{1}{g_1^2(x^1)} & 0 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \frac{1}{g_n^2(x^n)} \end{pmatrix}, \tag{D}$$

where the functions $\frac{1}{g_i}$ admit upper bounded primitives.

Remark 1.1. Such kind of metrics are used by Papa Quiroz [8] and Rapcsák [9], [10] to solve wide classes of problems arising on linear optimizations and nonlinear optimizations, respectively.

It is known [8] that this metric has as Christoffel coefficients $\Gamma^i_{ii} = -\frac{1}{g_i(x^i)} \cdot \frac{\partial g_i(x^i)}{\partial x^i}$, for all $i = \overline{1,n}$, and 0 otherwise. Moreover, $R^\ell_{ijk} = 0$, for all $i, j, k = \overline{1,n}$.

Remark 1.2. The metrics of diagonal type are particular cases of Hessian type metrics. Indeed, the decomposable function $H = \sum_{i=1}^n H_i(x^i)$, satisfies the following equations

$$\frac{\partial^2 H}{\partial x^i \partial x^j} = H_i''(x^i)\delta_{ij}, \quad i = \overline{1,n},\ j = \overline{1,n}.$$

The Hessian type metrics are useful tools in solving specific problems of WDVV (Witten-Dijkgraaf-Verlinde-Verlinde) equations of string theory [1]

2 Main results

Given (M, g) a Riemannian manifold, we denote by ∇ the Levi-Civita connection induced by the metric g.

Consider a function $f\colon M \to \mathbb{R}$, defined on an open domain, as closed mapping, that is $\{(f(P), P),\ P \in \operatorname{dom}(f)\}$ is a closed set in the product manifold $\mathbb{R} \times M$. Suppose f be at least three times differentiable.

Definition 2.1. The function f is said to be *k-self-concordant*, $k \geq 0$, with respect to the Levi-Civita connection ∇ defined on M if the following condition holds:

$$\left|\nabla^3 f(x)(X_x, X_x, X_x)\right| \leq 2k \left(\nabla^2 f(x)(X_x, X_x)\right)^{\frac{3}{2}}, \quad \forall x \in M,\ \forall X_x \in T_x M.$$

We are looking for decomposable self-concordant functions $f\colon \mathbb{R}^n_+ \to \mathbb{R}$, of the form

$$f(x^1, x^2, \dots, x^n) = f_1(x^1) + f_2(x^2) + \cdots + f_n(x^n), \tag{2.1}$$

where $f_i\colon \mathbb{R}_+ \to \mathbb{R}$ are differentiable functions.

Remark 2.2. The form (2.1) is suggested by the linearity of the set of self-concordant functions [6].

It follows

$$\frac{\partial f}{\partial x^i} = \frac{\partial f_i}{\partial x^i}; \qquad \frac{\partial^2 f}{\partial (x^i)^2} = \frac{\partial^2 f_i}{\partial (x^i)^2}; \qquad \frac{\partial^2 f}{\partial x^i \partial x^j} = 0, \quad \forall i \neq j.$$

By direct calculation, we obtain $f_{,ij} = 0$, for all $i \neq j$; $f_{,ii} = \dfrac{\partial^2 f_i}{\partial (x^i)^2} + \dfrac{\frac{\partial g_i}{\partial x^i}}{g_i(x^i)} \cdot \dfrac{\partial f_i}{\partial x^i}$; $f_{,ijk} = 0$ if at least two of the three indices i, j, k are different, and

$$f_{,iii} = \frac{\frac{\partial}{\partial x^i}\left[g_i(x^i)\frac{\partial}{\partial x^i}\left(g_i(x^i)\frac{\partial f_i}{\partial x^i}\right)\right]}{g_i^2(x^i)}.$$

If we put

$$f_i'' = \frac{\partial^2 f_i}{\partial (x^i)^2}; \qquad f_i' = \frac{\partial f_i}{\partial x^i}, \quad g_i' = \frac{\partial g_i}{\partial x^i},$$

then the covariant derivatives of the second order and of the third order of the function f have the forms:

$$(\nabla^2 f)(X, X) = \sum_{i=1}^n f_{,ii}(X^i)^2 = \sum_{i=1}^n \left[f_i''(x^i) + \frac{g_i'(x^i)}{g_i(x^i)} f_i'(x^i)\right](X^i)^2$$

and

$$(\nabla^3 f)(X, X, X) = \sum_{i=1}^n f_{,iii}(X^i)^3 = \sum_{i=1}^n \frac{\left[g_i(x^i)\left(f_i'(x^i)g_i(x^i)\right)'\right]'}{g_i^2(x^i)}(X^i)^3.$$

According to Definition 2.1, the condition for f to be self-concordant is

(2.2)

$$\left[\sum_{i=1}^n \frac{\left[g_i(x^i)\left(f_i'(x^i)g_i(x^i)\right)'\right]'}{g_i^2(x^i)}(X^i)^3\right]^2 \leq 4k^2\left[\sum_{i=1}^n \left(f_i''(x^i) + \frac{g_i'(x^i)}{g_i(x^i)} f_i'(x^i)\right)(X^i)^2\right]^3,$$

for all $x^i \in \mathbb{R}_+$ and all $X^i \in \mathbb{R}$.

If we use the relation

$$f_i''(x^i) + \frac{g_i'(x^i)}{g_i(x^i)} f_i'(x^i) = \frac{g_i(x^i)\left(f_i'(x^i)g_i(x^i)\right)'}{g_i^2(x^i)}$$

and introduce

$$F_i(x^i) = g_i(x^i)\left(f_i'(x^i)g_i(x^i)\right)', \tag{2.3}$$

the inequality (2.2) can be written as

$$\left[\sum_{i=1}^{n} \frac{F_i'(x^i)}{g_i^2(x^i)}(X^i)^3\right]^2 \leq 4k^2 \left[\sum_{i=1}^{n} \frac{F_i(x^i)}{g_i^2(x^i)}(X^i)^2\right]^3.$$

In the following, we need

Lemma 2.3. *If a_i and b_i are real numbers, and $b_i \neq 0$, $i = \overline{1,n}$, then*

$$\left(\sum_{i=1}^{n} \frac{a_i^3}{b_i^3}\right)^2 \leq \left(\sum_{i=1}^{n} \frac{a_i^2}{b_i^2}\right)^3.$$

The *proof* is a consequence of the Cauchy-Schwarz inequality ∎

Using Lemma 2.3, we have

$$\left[\sum_{i=1}^{n} \frac{F_i'(x^i)}{g_i^2(x^i)}(X^i)^3\right]^2 = \left[\sum_{i=1}^{n} \frac{(X^i)^3}{\left(\dfrac{g_i(x^i)}{\sqrt[3]{F_i'(x^i)g_i(x^i)}}\right)^3}\right]^2 \leq \left[\sum_{i=1}^{n} \frac{(X^i)^2}{\left(\dfrac{g_i(x^i)}{\sqrt[3]{F_i'(x^i)g_i(x^i)}}\right)^2}\right]^3$$

$$= \left[\sum_{i=1}^{n} \left(\sqrt[3]{F_i'(x^i)g_i(x^i)}\right)^2 \cdot \frac{(X^i)^2}{g_i^2(x^i)}\right]^3. \tag{2.4}$$

But f must verify the inequality (2.2). In this respect, we constrain the right side of (2.4) to be less than or equal to the right side of the inequality (2.2):

$$\left[\sum_{i=1}^{n} \left(\sqrt[3]{F_i'(x^i)g_i(x^i)}\right)^2 \cdot \frac{(X^i)^2}{g_i^2(x^i)}\right]^3 \leq \left[\sum_{i=1}^{n} \sqrt[3]{4k^2} \cdot F_i(x^i) \cdot \frac{(X^i)^2}{g_i^2(x^i)}\right]^3.$$

Theorem 2.4. *Suppose the function F is defined as in (2.3). If the connection ∇ is generated by g, then sufficient conditions for the function f to be self-concordant with respect to ∇ are given by*

$$F_i(x^i) \geq 0, \quad \left(\sqrt[3]{F_i'(x^i)g_i(x^i)}\right)^2 \leq \sqrt[3]{4k^2} \cdot F_i(x^i), \quad \forall i = \overline{1,n}.$$

Remark 2.5. The sufficient conditions in Theorem 2.4 imply the study of two cases as in the following. On one hand, we have to study the case of a differential equality, and on the other hand the case of a differential inequality.

CASE I OF DIFFERENTIAL EQUALITY.

In this case, by $\displaystyle\int \frac{1}{g_i(x^i)}\,dx^i$ we mean a negative primitive of the function $\dfrac{1}{g_i}$, that is $\displaystyle\int \frac{1}{g_i(x^i)}\,dx^i < 0$.

Let us determine the class of k-self-concordant functions f such that

$$\left(\sqrt[3]{F_i'(x^i)g_i(x^i)}\right)^2 = \sqrt[3]{4k^2} \cdot F_i(x^i), \quad \forall i = \overline{1,n}.$$

We use the $\frac{3}{2}$ power and we integrate. It follows

$$\left(F_i(x^i)\right)^{\frac{1}{2}} = -\frac{1}{k \displaystyle\int \frac{1}{g_i(x^i)} dx^i}.$$

But the right side must be non-negative and $k > 0$. We obtain $\displaystyle\int \frac{1}{g_i(x^i)}\, dx^i < 0$ and

$$F_i(x^i) = \frac{1}{k^2 \left(\displaystyle\int \frac{1}{g_i(x^i)} dx^i\right)^2}. \tag{2.5}$$

Taking into account the two forms of F_i given in (2.3) and (2.5), by integration, we find

$$f_i(x^i) = \frac{1}{k^2} \int \left[\frac{1}{g_i(x^i)} \cdot \int \frac{1}{g_i(x^i) \left(\displaystyle\int \frac{1}{g_i(x^i)} dx^i\right)^2} dx^i\right] dx^i. \tag{2.6}$$

Therefore, we proved

Theorem 2.6. *Let us suppose that the manifold $M = \mathbb{R}^n_+$ is endowed with the diagonal metric* (D)*, where the functions g_i satisfy the inequalities $\displaystyle\int \frac{1}{g_i(x^i)} dx^i < 0$, for all $i = \overline{1,n}$. If the functions f_i, $i = \overline{1,n}$, are given by* (2.6)*, then the decomposable function f, defined by* (2.1)*, is k-self-concordant.*

Examples:

1. Let $M = \mathbb{R}^n_+$ and $g_i(x^i) = e^{x^i}$. Then $\displaystyle\int \frac{1}{g_i(x^i)} dx^i = -e^{-x^i} < 0$.

Therefore, we can use Theorem 2.6 and we find a k-self-concordant function defined by

$$f\colon \mathbb{R}^n_+ \to \mathbb{R}, \quad f(x^1, x^2, \ldots, x^n) = \frac{1}{k^2}(x^1 + x^2 + \cdots + x^n).$$

2. Let $M = \mathbb{R}^n_+$ and $g_i(x^i) = -\frac{1}{x^i}$. Then $\displaystyle\int \frac{1}{g_i(x^i)} dx^i = -\frac{(x^i)^2}{2} < 0$.

Therefore, we can use Theorem 2.6 and we find a k-self-concordant function defined by

$$f\colon \mathbb{R}^n_+ \to \mathbb{R}, \quad f(x^1, x^2, \ldots, x^n) = -\frac{2}{k^2}(\ln x^1 + \ln x^2 + \cdots + \ln x^n).$$

Remark 2.7. To make a computer aided study of k-self-concordant functions we can perform symbolic computations for integrals. In this respect, we recommend the MAPLE software package [2], [13].

Case II of differential inequality.

We determine a class of decomposable self-concordant functions satisfying the differential inequality

$$\left(\sqrt[3]{F'(t)g(t)}\right)^2 \leq \sqrt[3]{4k^2}\cdot F(t)$$

and with given initial conditions $F(a)$ and $F'(a)$, $a > 0$ when $g(t) > 0$, for all $t > 0$. In this respect, we shall use

Lemma 2.8. Let the functions F and H of C^1-class defined on $[a, \infty)$ be given. If $F(a) = H(a)$ and $F'(t) \leq H'(t)$, for all $t \geq a$, then $F(t) \leq H(t)$, for all $t \geq a$.

The inequality given above can be written as

$$\frac{F'(t)}{F(t)^{\frac{3}{2}}} \leq \frac{2k}{g(t)}, \quad t \geq a. \tag{2.7}$$

The function

$$H(t) = \frac{1}{\left(k\displaystyle\int_a^t \frac{1}{g(s)}ds + \frac{1}{\sqrt{F(a)}}\right)^2},$$

satisfies the conditions $H'(t) = \dfrac{2k}{g(t)}H^{\frac{3}{2}}(t)$ and $H(a) = F(a)$.

We remark that the inequality (2.7) can be written as $\dfrac{F'(t)}{F(t)^{\frac{3}{2}}} \leq \dfrac{H'(t)}{H(t)^{\frac{3}{2}}}$, $t \geq a$, and by Lemma 2.8, $F(t) \leq H(t)$, for all $t \geq a$. Therefore

$$F(t) \leq \frac{1}{\left(k\displaystyle\int_a^t \frac{1}{g(s)}ds + \frac{1}{\sqrt{F(a)}}\right)^2}, \quad t \geq a > 0.$$

Since $F(t) = g(t)(f'(t)g(t))'$, we have

$$(f'(t)g(t))' \leq \frac{1}{g(t)} \cdot \frac{1}{\left(k\displaystyle\int_a^t \frac{1}{g(s)}ds + \frac{1}{\sqrt{F(a)}}\right)^2}.$$

If we integrate on the interval $[a, t]$, we obtain

$$f'(t)g(t) - f'(a)g(a) \leq \int_a^t \frac{1}{g(s)} \cdot \frac{1}{\left(k\displaystyle\int_a^s \frac{1}{g(\sigma)}d\sigma + \frac{1}{\sqrt{F(a)}}\right)^2} ds.$$

Then

$$f'(t) \leq \frac{1}{g(t)}\left[\frac{1}{\left(k\displaystyle\int_a^s \frac{1}{g(\tau)}d\tau + \frac{1}{\sqrt{F(a)}}\right)^2} ds + f'(a)g(a)\right].$$

If we integrate once again, we get

$$f(t) \leq \int_a^t \frac{1}{g(\tau)} \left[\int_a^\tau \frac{1}{\left(k \int_a^s \frac{1}{g(\tau)} d\tau + \frac{1}{\sqrt{F(a)}} \right)^2} ds + f'(a)g(a) \right] d\tau + f(a).$$

Theorem 2.9. *Let us suppose that the manifold $M = \mathbb{R}^n_+$ is endowed with the diagonal metric* (D)*. If the functions f_i, $i = \overline{1,n}$, are given by*

$$f_i(x_i) \leq \int_a^{x_i} \frac{1}{g_i(\tau)} \left[\int_a^\tau \frac{1}{\left(k \int_a^s \frac{1}{g_i(\tau)} d\tau + \frac{1}{\sqrt{F_i(a)}} \right)^2} ds + f_i'(a)g_i(a) \right] d\tau + f_i(a),$$

and g_i are positive functions for all $i = \overline{1,n}$, then the decomposable function f, defined by (2.1)*, is k-self-concordant.*

We can change the point of view. We can ask to find decomposable functions f which are both self-concordant and generate the metric g, that is we have $\frac{1}{g_i} = f_i''$, for all $i = \overline{1,n}$. Using (2.6), we find

Theorem 2.10. *The Shanon entropy [11] function*

$$f\colon \mathbb{R}^n_+ \to \mathbb{R}, \quad f(x^1, x^2, \ldots, x^n) = \frac{1}{k^2}(\ln k^2x^1 + \ln k^2x^2 + \cdots + \ln k^2x^n),$$

is self self-concordant.

Open problem. Find other types of self-concordant functions with respect to metrics of diagonal type.

Acknowledgements

The authors would like to thank to professors Constantin Udrişte and Ionel Ţevy for their fruitful comments on the preliminary version of the paper.

References

[1] H. W. Braden and A. Marshakov, *WDVV equations as functional relations*, arXiv:hep-th/0205308 v1, May 2002.

[2] Maria Teresa Calapso and C. Udrişte, *Isothermic surfaces as solutions of Calapso PDE*, Balkan J. Geom. Appl., 13, 1 (2008), 20-26.

[3] V. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Springer-Verlag, London, 1994.

[4] D. den Hertog, *Interior point approach to linear, quadratic and convex programming*, Mathematics and Its Applications (277), Kluwer, 1994.

[5] D. Jiang, J. B. Moore and H. Ji, *Self-concordant functions for optimization on smooth manifolds*, 43rd IEEE Conf. Decision and Control, Atlantis 2004, 3631-3636.

[6] Y. Nesterov and A. Nemirovsky, *Interior-point polynomial algorithms in convex programming*, Studies in Applied Mathematics (13), Philadelphia, 1994.

[7] Y. Nesterov and M. J. Todd, *On the Riemannian geometry defined by self-concordant barriers and interior point methods*, Found. Comp. Math. 2, 4(2002), 333-361.

[8] E. A. Quiroz and P. R. Oliveira, *New results on linear optimization through diagonal metrics and Riemannian geometry tools*, Technical Report ES-654/04, PESC COPPE, Federal University of Rio de Janeiro, 2004.

[9] T. Rapcsák, *Smooth Nonlinear Optimization in* $\mathbb{R}^n$, Kluwer Academic Publishers, 1997.

[10] T. Rapcsák, *Geodesic convexity in nonlinear optimization*, JOTA 69, 1 (1991), 169-183.

[11] T. Schürmann, *Bias analysis in entropy estimation*, J. Phys. A: Math. Gen. 37 (2004), L295-L301.

[12] C. Udrişte, *Optimization methods on Riemannian manifolds*, Algebra, Groups and Geometries, 14 (1997), 339-359.

[13] C. Udrişte, *Tzitzeica theory - opportunity for reflection in Mathematics*, Balkan J. Geom. Appl. 10, 1 (2005), 110-120.

Authors' addresses:

Gabriel Bercu: University "Dunărea de Jos" of Galaţi, Department of Mathematics, Domnească Street, No. 47, 800008 Galaţi, Romania, gbercu@ugal.ro

Mihai Postolache: University "Politehnica" of Bucharest, Faculty of Applied Sciences Splaiul Independenţei, No. 313, 060042 Bucharest, Romania, mihai@mathem.pub.ro