Aust. J. Math. Anal. Appl.
Vol. **17** (2020), No. 2, Art. 14, 12 pp.
AJMAA

OPTIMIZATION TECHNIQUES ON AFFINE DIFFERENTIAL MANIFOLDS

ALI S RASHEED[1], FAIK MAYAH[2] AND AHMED A H AL-JUMAILI[1]

Received 28 March, 2020; accepted 10 September, 2020; published 23 November, 2020.

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH, IRAQ.[1]
ali.math2018@yahoo.com, ahmedhashem@gmail.com

DEPARTMENT OF PHYSICS, COLLEGE OF SCIENCES, UNIVERSITY OF WASIT, IRAQ.[2]
faik.mayah@gmail.com

ABSTRACT. In addition to solid ground of Riemannian manifolds fundamentals, this article interviews some popular optimization methods on Riemannian manifolds. Several optimization problems can be better stated on manifolds rather than Euclidean space, such as interior point methods, which in turns based on self-concordant functions (logarithmic barrier functions). Optimization schemes like the steepest descent scheme, the Newton scheme, and others can be extended to Riemannian manifolds. This paper introduces some Riemannian and non-Riemannian schemes on manifolds.

Key words and phrases: Convex programming on affine manifolds; Barrier function; Special connections; Complexity analysis; Affine convex functions.

2010 *Mathematics Subject Classification.* 90C25, 90C51.

ISSN (electronic): 1449-5910

1. INTRODUCTION

Recently, optimization on Riemannian manifolds has become an interesting ongoing research due to its positives over Euclidean space. The notion of this is that Riemannian optimization depends on the curved manifolds, thereby constraints can be eliminated and only feasible points required. Riemannian optimization has several applications in diverse sectors. For instance, machine learning, computer vision and data mining, Riemannian dictionary learning, and tensor clustering, see respectively, [8]. Extensive research to solve minimizing problems on Riemannian manifolds available in [2],[40]. The constrained optimization algorithms on a manifold have lower complexity, dimension as well as better numerical properties. Besides, classical optimization schemes like the steepest descent scheme, the Newton scheme, and others can be extended to Riemannian manifolds; see, e.g., [19], [37], [3]. Nowadays, solid Riemannian optimization techniques have become available on http://www.manopt.org.
In the last two decades, Interior Point Methods (IPM) have most extremely studied methods for both linear and certain convex programming problems, due to their excellent computations and properties; see, e.g., [23], [31], [43], [44]. Chronologically, conventional methods such as the steepest descent method, conjugate gradient method or Newton method have been used widely to minimize a cost function [22]. However, several inequality constraint optimization problems have been raised in trade and industry resulted in developing new optimization techniques; see, e.g., [30], [42]. The basic concept of the IPM is to transform a constrained problem into a parameterized unconstrained one using a barrier function, known as a self-concordant function (S-CF) defined by [31]. The keyword identified by [13] is the construction of so-called S-CF for the constraint set (neglecting the equality constraints). The significance of these functions is demonstrated into two points. Firstly, they connected with logarithmic barrier functions in IPM. Secondly, the proposed damped Newton method for optimizing S-CF, which involve known parameters. However, most IPM forms based on eliminating the constraint set and adding a multiple of the barrier function. The central path is a result of set of minimizers work as the multiplier changes for the problem. This approach applicable to obtain so-called primal central path or primal-dual central path based on choice of either the original problem or the original problem and its dual simultaneously, respectively. According to [32], the corresponding algorithm takes a sequence of short steps. Thus, it's reasonable to consider the corresponding Riemannian metric defined on the interior of the constraint set. Moreover, we can consider the shortest paths in this metric, or geodesic curves. Affine manifolds are smooth manifolds which are locally modeled on a finite dimensional affine vector space. These manifolds arise naturally in differential geometry, theoretical physics and optimization theory. While much more is known about their Riemannian counter parts, there are several outstanding conjectures regarding these manifolds.

2. RIEMANNIAN GEOMETRY PRELIMINARIES

The solid ground to understand the definition of Riemannian manifolds requires definition of several concepts such as smooth manifolds, charts, atlases and tangent vectors seen as equivalence classes of the exposition of these definitions is mainly inspired from [2], [11], [12], [24], [33], [9].

2.1. **Charts and Manifolds.** Manifolds are sets, which locally identified with patches of R^m. A set of these identifications forms so-called charts. Then, atlas is a set of compatible charts that covers the whole set. Both the atlas and the set constitute a manifold.

Definition 2.1. (chart). Let υ be a set, $U \subset \upsilon$, and φ is a bijection between U and an open set of R^m . A chart of υ is a pair (U, φ). U is the chart's domain and m is the chart's dimension. Let $q \in U$, then elements of $\varphi(q) = (x_1, ..., x_m)$ are called the coordinates of q in the chart (U, φ).

Definition 2.2. (compatible charts)

Let $(U, \varphi), (W, \psi)$ be two charts of υ, of dimensions n and m, respectively, are smoothly compatible (C^∞ - compatible) if either $U \bigcap W = \varphi$ or $U \bigcap W \neq \phi$ and $\mathbf{.}\varphi(U, W)$ is an open set of R^n, $\mathbf{.}\psi(U, W)$ is an open set of R^m, $\mathbf{.}\psi \circ \varphi^{-1} : \varphi(U, W) \rightarrow \psi(U, W)$ is a smooth diffeomorphism. When $U \bigcap W \neq \phi$, implies n = m.

Definition 2.3. (atlas). A set $B = (U_i, \varphi_i), i \in I$ of pairwise smoothly compatible charts: $\bigcup_{i \in I} U_i = \upsilon$ is a smooth atlas of υ. Two atlases B_1 and B_2 are compatible if $B_1 \cup B_2$ is an atlas. Atlas B and all the charts compatible with B called a unique maximal atlas B^+.

Definition 2.4. (manifold).

A pair $\mu = (\upsilon, B^+)$ defines a smooth manifold where υ is a set and B^+ is a maximal atlas of υ.

Definition 2.5. (dimension).

Given a manifold $\mu = (\upsilon, B^+)$, if the all charts of B^+ have the same dimension m, then m is the dimension of μ.

Definition 2.6. (smooth mapping). Given μ,$\aleph$, two smooth manifolds. A mapping $g : \mu \rightarrow \aleph$ is of class C^r if, for all q in υ, there is chart (U, φ) of υ and a chart (W, ψ) of $\aleph$ such that $q \in U, g(U) \subset W$ and $\psi \circ g \circ \varphi^{-1} : \varphi(U) \rightarrow \psi(W)$ is of class C^r. This is called the local expression of g in the charts (U, φ) and (W, ψ). A smooth map is of class C^∞.
More definitions related to tangent spaces and tangent vectors are available in [9], page 24.

2.2. **Riemannian structure and gradients.**

Definition 2.7. (inner product). Given υ a smooth manifold and fix point $g \in \upsilon$. Thus, An inner produc $\langle \cdot, \cdot \rangle q$ on a tangent space $T_q\upsilon$ is a bilinear, symmetric positive definite form on $T_q\upsilon$. Mathematically, $\forall X, Y, Z \in T_q\upsilon, \alpha, \beta \in R$:
$\mathbf{.}\langle X, X \rangle q \geqslant 0, and \langle X, X \rangle q = 0 \Leftrightarrow X = 0$,
$\mathbf{.}\langle \alpha X + \beta Y, Z \rangle q = \alpha \langle X, Z \rangle q + \beta \langle Y, Z \rangle q$, and
$\mathbf{.}\langle X, Y \rangle q = \alpha \langle Y, X \rangle q$.

Let υ be a smooth manifold and h is a Riemannian metric. Then, the pair (υ, h) called a Riemannian manifold. A Riemannian metric is a smoothly varying inner product defined on the tangent spaces of υ. This means that for each $q \in \upsilon, h_q(\cdot, \cdot) = \langle \cdot, \cdot \rangle q$ is an inner product on $T_q\upsilon$.

Definition 2.8. (gradient).

The gradient of a scalar field g on a Riemannian manifold υ at point g, obtained by grad $g(q)$ as the unique element of $T_q\upsilon$ satisfying:

$$Dg(q)[X] = \langle \text{grad } g(q), X \rangle q, \forall X \in T_q\upsilon.$$

Then, grad $g : \upsilon \rightarrow T_\upsilon$ represents a vector field on υ.

2.2.1. *Riemannian submanifolds.* Let (υ, f) be a Riemannian manifold, and let $\bar{\upsilon}$ be a submanifold of υ and $\bar{f}$ is the restriction of f to the tangent spaces of $\bar{\upsilon}$. Then, $(\bar{\upsilon}, \bar{f})$ called a Riemannian submanifold. This means that $\forall q \in \bar{\upsilon}$ and $\forall X, Y \in T_q\bar{\upsilon} \subset T_q\upsilon$, the metrics f and f are compatible in the sense that $f_q(X, Y) = \bar{f}_q(X, Y)$.

2.2.2. *Riemannian quotient manifolds.* Let (υ, f) be a Riemannian manifold, and let $\bar{\upsilon} = \upsilon / \sim$ be a quotient manifold of υ. We will now leverage the Riemannian structure of υ to equip $\bar{\upsilon}$ with a Riemannian structure as well. Therefore, we first single out one horizontal distribution as follows.

$$H_q := U_q^\perp = X \in T_q\upsilon : f_q(X, Y) = 0, \forall Y \in U_q, \forall q \in \upsilon.$$

2.3. **Distances and geodesic curves.** A line segment on R^n is defined as curves with arc-length parameterization is that they have zero acceleration. First of all, let's introduce the concept of tangent vectors to curves or so-called velocity vectors. Consider Γ a curve of class C^1 such that $\Gamma : [a, b] \to \upsilon$, and $t \in [a, b]$, define another such curve on υ by shifting its parameter:

$$\Gamma_t : [a - t, b - t] \to \upsilon : \tau \to \Gamma_t(\tau) = \Gamma(t + \tau).$$

Thus, the equivalence class $[\Gamma_t] \in T\Gamma(t)\upsilon$ is a vector tangent to Γ at time t, $\langle X, Y\rangle_q \triangleq$. For more understanding, we have to see the next definitions.

Definition 2.9. (acceleration along a curve).

Given υ a smooth manifold equipped with a connection ∇. Let I be an open interval of R is a C^2 curve on υ. The acceleration along Γ is obtained by:

$$t \to \nabla_{\dot{\Gamma}(t)}\dot{\Gamma}(t) \in T_{\Gamma(t)}\upsilon.$$

Definition 2.10. (geodesic).

Let I be an open interval of **R**, and a curve $\Gamma : I \to \upsilon$. Thus, υ is geodesic if and only if it has zero acceleration on all its domain.

Definition 2.11. (length of a curve). The length of a curve of class C^1, $\Gamma : [a, b] \to \upsilon$, on a Riemannian manifold (υ, f), with $\Gamma(t)^{\cdot} \triangleq f_q(X, Y)$, is defined by

$$\text{Length}(\Gamma) = \int_\alpha^\beta \sqrt{\langle\dot{\Gamma}(t), \dot{\Gamma}(t)\rangle_{\Gamma(t)dt)}} = \int_\alpha^\beta ||\dot{\Gamma}(t)||_{\Gamma(t)}dt.$$

Definition 2.12. (Riemannian distance).

It's called sometimes (geodesic distance). Then, it's given on υ by $dist$:

$$\upsilon \times \upsilon \to \mathbb{R}^+ : (p, q) \mapsto dist(p, q) = \inf_{\Gamma \in \jmath} \; Lengh(\Gamma),$$

where is the set of all C^1 curves $\Gamma : [0, 1] \to \upsilon : \Gamma(0) = p$, and $\Gamma(1) = q$.

2.4. **Parallel translation.** It's difficult to compare vectors from different tangent spaces on manifolds. As a result parallel translation comes to transport vectors between tangent spaces without lost the information. The idea, vector transport defines how to transport a vector $X \in T_q\upsilon$ from a point $q \in \upsilon$ to a point $R_q(Y) \in \upsilon, Y \in T_q\upsilon$. However, let's introduce the Whitney sum then quote the definition of vector transport.

$$T_\upsilon \oplus T_\upsilon = (X, Y) : Y, X \in T_q\upsilon, q \in \upsilon$$

Thereby,$T_\upsilon \oplus T_\upsilon$ is the set of pairs of tangent vectors belonging to a same tangent space.

Definition 2.13. (vector transport). A vector transport on a manifold υ is a smooth mapping

$$\text{Transp} : T_\upsilon \oplus T_\upsilon \to T_\upsilon : (X, Y) \mapsto Transp_Y(X),$$

satisfies $\forall q \in \upsilon$

(1) There exists a retraction R associated with Transp, such that $Transp_Y(X) \in (X) \in T_Rq(Y)\upsilon$,
(2) $Transp_0(X) = X, \forall X \in T_q\upsilon$, (consistency),
(3) $Transp_Y(aX+bY) = aTransp_Y(X)+bTransp_Y(Z), \forall X, Y, Z \in T_q\upsilon \in$ R(Linearity)

2.5. **Curvature.**

Definition 2.14. (Riemannian curvature tensor). Given υ a Riemannian manifold, and ∇ a Riemannian connection, then the Riemannian curvature tensor $R : \chi(\upsilon) \times \chi(\upsilon) \times (\upsilon) \to \chi(\upsilon)$ is obtained by

$$\Re(\zeta\eta)\xi = \nabla_\zeta \nabla_\eta \xi - \nabla_\eta \nabla_\zeta \xi - \nabla_{[\zeta,\eta]}\xi.$$

The Riemannian manifold is υ if and only if $\Re$ vanishes identically.

3. OPTIMIZATION ON SMOOTH MANIFOLDS

Several optimization problems can be better stated on manifolds rather than Euclidean space, such as interior point methods, which in turns based on self-concordant functions (logarithmic barrier functions), see([19], [20], [34]). Therefore, many optimization problems are extended to non-Euclidean spaces; see, e.g., [37], [3], [20]. For instance, finding the largest eigenvalue of a symmetric matrix, and Optimization problems subject to nonlinear differentiable equality constraints on Euclidean space. Such as these examples are endowed with a Riemannian metric [37]. Moreover, [32] have revealed that a Riemannian metric can be represented by a self-concordant barrier function. Thereby a metric gives a good clarification of the optimal direction for algorithms. Further up, the self-concordant concept has been applied in [39] to present a gentle interpretation of a logarithm cost function optimized on a manifold. This study extended by [35]. Besides, optimization schemes like the steepest descent scheme, the Newton scheme, and others can be extended to Riemannian manifolds; see, e.g., [19], [37], [3]. This is in turn raising a question, what are results of extending the meaning of Riemannian metric to affine connections on Riemannian manifolds? However, for more understanding, this article reviews some Riemannian and non-Riemannian approaches.

3.1. **Riemannian Schemes.**

3.1.1. *Steepest descent method on manifolds ($SDMOM$).* $SDMOM$ is the simplest classical algorithm for the optimization on Riemannian manifolds, introduced by [25], [26] and [18]. Unfortunately, $SDMOM$ has a slow linear convergence rate. On the other hand in 90s, this method applied to solve problems in the control theory; see, e.g., [19], [20], [10], [38], [27].

3.1.2. *Newton method on manifolds ($NMOM$).* Compare to the $SDMOM$, the $NMOM$ has faster quadratic convergence rate. It has been extended to a Riemannian sub-manifold of R^n by (Gabay,1982). Further progress of the $NMOM$ on Riemannian manifolds have proposed by [40], [37], [27], [28], [14], [13].

3.1.3. *Quasi-Newton method on manifolds ($QNMOM$).* To overcome the computational cost of the $NMOM$ due to inverse of a symmetric matrix calculation, the $QNMOM$ has been presented by [37]. This method uses only the first order information of the cost function to approximate the Hessian inverse and has a super-linear local convergence rate. As a result, researchers developed diverse algorithm forms of the $QNMOM$. For example, Davidon-Fletcher-Powell (DFP)[15] method and the Broyden [6], [5] Fletcher [16] Goldfarb [42] Shanno [43] ($BFGS$) method. In the early 80s, [18] firstly generalized the $BFGS$ method to a Riemannian manifold. Further up, [7] improved the $BFGS$ method on the Grassmann manifold and reduced the computational cost.

3.1.4. *Trust-region method (TRM).* To avoid the inverse Hessian matrix calculation with high convergence rate, [1] proposed TRM. One of TRM applications, for instance, is Grassmann manifold for the matrix completion problem [8]. Practically, each iteration of TRM involves solving the Riemannian Newton equation [38], which leads to increase the algorithm complexity. This is resulted in developing the method by [40] without solving Riemannian Newton equation.

3.2. **Non-Riemannian Schemes.** The classical optimization methods on a manifold are endowing with a metric structure to be into a Riemannian manifold, see [29]. For example, the gradient is replaced by the Riemannian gradient. To minimize a cost function, [30] described a general context for developing numerical algorithms on a manifold. Based on this development, the corresponding steepest descent and Newton methods reduced the computational complexity compared with the Riemannian methods.

4. SELF-CONCORDANT FUNCTIONS ON RIEMANNIAN MANIFOLDS (S-CF ON RM)

The S-CF play a vital role in solving certain convex constrained optimization problems. Its extension to Riemannian manifolds has been derived by [39]. This author deliberated the following convex programming problem

$$\min[g_\circ(x) : g_i(x) \leqslant 0, i = 1, 2, , n, x \in \upsilon] \tag{4.1}$$

Where υ is a complete n-dimensional Riemannian manifold. Interior point methods associated with barrier functions were used for solving. Besides, this barrier function is selected to be self-concordant.

Assumption 1.
Let υ be a geodesic complete smooth manifold of finite dimension and let ∇ be a symmetric affine connection on it. In addition, the geodesic between two points is unique, then.

$$g : \upsilon \to R, \tag{4.2}$$

$$\nabla_W \nabla_Q(g) = \nabla_Q \nabla_W(g), \tag{4.3}$$

For any vector field W , $Q.g$ has an open domain and closed map.

Definition 4.1. g is a self-concordant function respect to

$$\nabla \iff |\nabla^3_W g(q)| \leqslant \nu_g [\nabla^2_W g(q)]^{3/2}, \quad \forall W \in T_q\upsilon, \tag{4.4}$$

where υ is the tangent space of vat point q and υ_g is a positive constant. $\upsilon_g = 2$ can be selected as a constant according to [21].

Assumption 2

$$\nabla^2_W g(q) > 0, \quad \forall q \in dom(g), \quad W \in T_q\upsilon. \tag{4.5}$$

Thereby, we can define a Dikin-type ellipsoid $X^0(q; r)$ based on the second-order covariant differentials as follows.

Definition 4.2.

$$X^0(q;r) = (p \in \upsilon | [\nabla^2_{W_{qp}} g(q)]^{0.5} < r), \forall q \in dom(g), r > 0, \tag{4.6}$$

Such that $W_q p$ is the vector field defined by the geodesic connecting the points q and p. Basically, definition of S-CF depends on second order and third order covariant differentials with respect to the same vector field W. Furthermore, [21] in his work pointed out four gentle properties of the S-CF.

Assumption 3.
Let (υ, h) be a Riemannian manifold, we observe by ∇ the Levi-Civita connection induced by the metric h. Consider a mapping $g : \upsilon \to \mathbb{R}$, such that $[(g(q), q), q \in dom(g)]$ is a closed set in the product manifold $\mathbb{R}$.

Definition 4.3. The function g is called r-self-concordant, $r \geqslant 0$, with respect to the Levi-Civita connection ∇ defined on υ if:

$$|\nabla^3_g(y)(Y_y, Y_y, Y_y)| \leqslant 2r(\nabla^2_g(y)(Y_y, Y_y))^{3/2} \forall y \in \upsilon, \forall Y_y \in T_Y \upsilon.$$

Definition 4.4. (Newton decrement).
Let the auxiliary quadratic cost defined on $T_q\upsilon$ as follows.

$$\xi_g.q(W) := g(q) + \nabla_W g(q) + 0.5 \nabla^2_W g(q), \tag{4.7}$$

then,

$$W_\xi(g, q) := argmin_{W \in T_q \upsilon} \xi_g.q(W), \tag{4.8}$$

This means that the Newton decrement $W_\xi(g, q)$ is defined as the minimum solution to the $\xi\xi_g.q(W)$.

Theorem 4.1. *Consider the following S-CF $g : \upsilon \to R, q$ a given point in $dom(g) \subseteq \upsilon$, and W_ξ is the Newton decrement defined at q, thus*

$$\nabla_W \xi \nabla_W g(q) = -\nabla_W g(q), \forall W \in T_q \upsilon, \tag{4.9}$$

$$(\nabla_W \xi^2 g(q))^{0.5} = max[g(q) | W \in T_q \upsilon, \nabla^2_W g(q) \leqslant 1]. \tag{4.10}$$

4.0.1. *Damped Newton method for S-CF.* Damped Newton method (DNM) of S-CFs can be understood by the next theorem.

Theorem 4.2. [21]. *Consider the following*

$$\lambda_g(q) := max_{W \in T_q \upsilon} |\nabla_W g(q)| / (\nabla^2_W g(q))^{0.5} : q \in dom(g). \tag{4.11}$$

For some $q \in$ dom(g), if $\lambda_g(q) < 1$, thus there exists a unique point $q^*_g \in dom(g)$:

$$g(q^*_g) = min[g(q), q \in dom(g)].$$

The feasible algorithm of DNM is obtained by two main steps.

1. *Find a feasible point $q^0 \in dom(g)$.*
2. $q_i = exp_{q_i - 1}(1/1 + \lambda_g(q_i - 1) W_\xi)$,

where $exp_{q_{i-1}} t\ W_\xi$ is the exponential map of the Newton decrement at q_{i-1}.

5. AFFINE CONVEX FUNCTIONS

Definition 5.1. An affine differential manifold (M, Γ) is called auto-parallely complete if any auto-parallel $x(t)$ starting at $p \in M$ is defined for all values of the parameter $t \in \mathbb{R}$.

Theorem 5.1. [4] *Let M be a (Hausdorff, connected, smooth) compact n-manifold endowed with an affine connection Γ, and let $p \in M$. If the holonomy group $Hol_p(\Gamma)$ (regarded as a subgroup of the group $Gl(T_pM)$ of all the linear automorphisms of the tangent space T_pM) has compact closure, then (M, Γ) is auto-parallely complete.*

Let (M, Γ) be an auto-parallely complete affine differential manifold. For a C^2 function $f\colon M \to \mathbb{R}$, we define the tensor $\mathrm{Hess}_\Gamma f$ of components

$$(\mathrm{Hess}_\Gamma f)_{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j} - \Gamma^h_{ij} \frac{\partial f}{\partial x^h}.$$

Definition 5.2. A C^2 function $f\colon M \to \mathbb{R}$ is called:

(1) linear affine with respect to Γ if $\mathrm{Hess}_\Gamma f = 0$, throughout M;
(2) affine convex (convex with respect to Γ) if $\mathrm{Hess}_\Gamma f \succeq 0$ (positive semidefinite), throughout M, (see [41].

The function f is:

(1) linear affine if its restriction $f(x(t))$ on each auto-parallel curve $x(t)$ satisfies $f(x(t)) = at + b$, for some numbers a, b that may depend on $x(t)$;
(2) affine convex, if its restriction $f(x(t))$ is convex on each auto-parallel curve $x(t)$.

Example 5.1. Consider the bilevel programming problem

$$f(x,y) = (x-1)^2 + (y-1)^2, \quad \text{subject to} \quad g(x,y) = x - 50y = 500 = 0$$

(1) we find the first and the second partial derivatives

$$\frac{\partial f}{\partial x} = 2(x-1),\ \frac{\partial f}{\partial y} = 2(y-1),\ \frac{\partial^2 f}{\partial x^2} = 2,\ \frac{\partial^2 f}{\partial x \partial y} = 0,\ \frac{\partial^2 f}{\partial y^2} = 2$$

(2)

$$\frac{\partial f}{\partial x^i \partial y^i} - \Gamma^h_{ij} = f_{,ij}$$

$$\begin{aligned} f_{,11} &= 2 - \Gamma^1_{11}(2x-2) + \Gamma^2_{11}(2y-2) \\ f_{,12} &= 0 - \Gamma^1_{12}(2x-2) + \Gamma^2_{12} \\ f_{,22}&(2y-2) \\ &= 2 - \Gamma^1_{22}(2x-2) + \Gamma^2_{22}(2y-2) \end{aligned}$$

(3)

$$\begin{pmatrix} f_{,11} & f_{,12} \\ f_{,12} & f_{,22} \end{pmatrix} = \begin{pmatrix} 2 - \Gamma^1_{11}(2x-2) - \Gamma^2_{11}(2y-2) & -\Gamma^1_{12}(2x-2) - \Gamma^2_{12}(2y-2) \\ -\Gamma^1_{12}(2x-2) + \Gamma^2_{12}(2y-2) & 2 - \Gamma^1_{22}(2x-2) - \Gamma^2_{22}(2y-2) \end{pmatrix}$$

Positive semidefinite when $f_{,11} . f_{,22} - f^2_{,12} \geq 0$, $f_{,11} \geq 0$

(4) We must find connections

$$2 - \Gamma^1_{11}(2x-2) - \Gamma^2_{11}(2y-2) \geq 0$$

$$(2-\Gamma^1_{11}(2x-2)-\Gamma^2_{11}(2y-2)).(2-\Gamma^1_{22}(2x-2)-\Gamma^2_{22}(2y-2))-(-\Gamma^1_{12}(2x-2)-\Gamma^2_{12}(2y-2))^2 \geq 0$$

We prefer to introduce the slackness variables $\omega_1 \geq 0$, $\omega_2 \geq 0$, to transform it into an equality system

$$2 - \Gamma^1_{11}(2x-2) - \Gamma^2_{11}(2y-2) - \omega_1 = 0$$

$$(2-\Gamma^1_{11}(2x-2)-\Gamma^2_{11}(2y-2)).(2-\Gamma^1_{22}(2x-2)-\Gamma^2_{22}(2y-2))-(-\Gamma^1_{12}(2x-2)-\Gamma^2_{12}(2y-2))^2-\omega_2 = 0$$

We eliminate the critical points of f, if they exists. A possible connection has the components

$$\Gamma^1_{11} = \frac{1}{(2x-2)}[2 - \Gamma^2_{11}(2y-2) - \omega_1] \quad \text{where (2x-2)} \neq 0, \Gamma^1_{11} \text{ arbitrary}$$

$$\Gamma^2_{22} = \frac{1}{(2y-2)}[\frac{\omega_2 + (-\Gamma^1_{12}(2x-2) - \Gamma^2_{12}(2y-2))^2}{(2-\Gamma^1_{11}(2x-2)\Gamma^2_{11}(2y-2))} + (2x-2)\Gamma^1_{22} - 2]$$

where (2y-2) $\neq 0$, $(2 - \Gamma^1_{11}(2x-2) - \Gamma^2_{11}(2y-2)) \neq 0$ and $\Gamma^1_{12}, \Gamma^2_{12}, \Gamma^1_{22}$ are arbitrary given.

By the same way to show that the constraint function,

$$\frac{\partial g}{\partial x} = 1,\ \frac{\partial g}{\partial y} = -50,\ \frac{\partial^2 g}{\partial x^2} = 0,\ \frac{\partial^2 g}{\partial x \partial y} = 0,\ \frac{\partial^2 g}{\partial y^2} = 0$$

$$\frac{\partial^2 g}{\partial x^i \partial y^i} - \Gamma^h_{ij}\frac{\partial g}{\partial x^h}, g_{,ij}$$

$$g_{,11} = 0 - \Gamma^1_{11} - 50\Gamma^2_{11}$$

$$g_{,12} = 0 - \Gamma^1_{12} - 50\Gamma^2_{12}$$

$$g_{,22} = 0 - \Gamma^1_{22} - 50\Gamma^2_{22}$$

$$-\Gamma^1_{11} - 50\Gamma^2_{11} \geq 0$$

$$(-\Gamma^1_{11} - 50\Gamma^2_{11}).(-\Gamma^1_{22} - 50\Gamma^2_{22}) - -\Gamma^1_{12} - 50\Gamma^2_{12})^2 \geq 0$$

$$-\Gamma^1_{11} - 50\Gamma^2_{11} - \omega_1 = 0$$

$$(-\Gamma^1_{11} - 50\Gamma^2_{11}).(-\Gamma^1_{22} - 50\Gamma^2_{22}) - -\Gamma^1_{12} - 50\Gamma^2_{12})^2 - \omega_2 = 0$$

$\Gamma^1_{11} = (-50\Gamma^2_{11} - \omega_1)$, where Γ^2_{11} arbitrary

$$\Gamma^2_{22} = \frac{-1}{50}\left(\frac{\omega_2(-\Gamma^1_{12} - 50\Gamma^2_{12}x^2)^2}{(-\Gamma^1_{11} - 50\Gamma^2_{11})} + \Gamma^1_{22}\right)$$

where $(-\Gamma^1_{12} - 50\Gamma^2_{11}) \neq 0$ and $\Gamma^1_{12}, \Gamma^2_{12}, \Gamma^1_{22}$ are arbitrary given.

The following unconstrained function $L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$ is called the Lagrangian of f w.r.t g,and λ is called the Lagrange multiplier, where $\lambda > 0$,

$$L(x,y,\lambda) = (x-1)^2 + (y-1)^2 + \lambda(x - 50y + 500)$$

$$\frac{\partial L}{\partial x} = 2x - 2 + \lambda, \quad \frac{\partial L}{\partial y} = 2y - 2 - 50\lambda, \quad \frac{\partial L}{\partial \lambda} = x - 50y + 500,$$

$$\frac{\partial^2 L}{\partial x^2} = 2, \quad \frac{\partial^2 L}{\partial x \partial y} = 0, \quad \frac{\partial^2 L}{\partial y^2} = 2$$

$$\frac{\partial^2 L}{\partial x^i \partial y^i} - \Gamma^h_{ij}\frac{\partial L}{\partial x^h} = L_{,ij}$$

$$\begin{aligned}
L_{,11} &= 2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda)\\
L_{,12} &= 0 - \Gamma^1_{12}(2x - 2 + \lambda) - \Gamma^2_{12}(2y - 2 - 50\lambda)\\
L_{,22} &= 2 - \Gamma^1_{22}(2x - 2 + \lambda) - \Gamma^2_{22}(2y - 2 - 50\lambda)
\end{aligned}$$

Now to show that

$$2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda) \geq 0$$
$$[2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda)(2 - \Gamma^1_{22}(2x - 2 - 50\lambda) -$$
$$(-\Gamma^2_{22}(2y - 2 - 50\lambda) - (-\Gamma^1_{12}(2x - 2 - 50\lambda) - \Gamma^2_{12}(2y - 2 - 50\lambda))^2] \geq 0$$
$$2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda) - \omega_1 = 0$$
$$[2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda)(2 - \Gamma^1_{22}(2x - 2 - 50\lambda) -$$
$$(-\Gamma^2_{22}(2y - 2 - 50\lambda) - (-\Gamma^1_{12}(2x - 2 + \lambda) - \Gamma^2_{12}(2y - 2 - 50\lambda))^2] - \omega_2 = 0$$

$$\Gamma^2_{11} = \frac{1}{(2y - 2 - 50\lambda)}\left(2 - \Gamma^1_{11}(2x - 2 + \lambda) - \omega_1\right)$$

where $(2y - 2 - 50\lambda) \neq 0$ and Γ^1_{11} arbitrary given.

$$\Gamma^1_{22} = \frac{-1}{(2x - 2 + \lambda)}\left[\frac{\omega_2 + (-\Gamma^1_{12}(2x - 2 + \lambda) - \Gamma^2_{12}(2y - 2 - 50\lambda))^2}{(2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda))} + \right.$$
$$\left. +(2y - 2 - 50\lambda)\Gamma^2_{22}\right],$$

where $(2x - 2 + \lambda) \neq 0$,

$$(2 - \Gamma^1_{11}(2x - 2 + \lambda) - \Gamma^2_{11}(2y - 2 - 50\lambda) \neq 0$$

and $\Gamma^1_{12}, \Gamma^2_{12}, \Gamma^2_{22}$ are arbitrary given.

6. CONCLUSION

It can be seen that several optimization methods have been reported in this article. Starting with the basic concepts of the Riemannian manifolds to understand the fundamentals up to self-concordant functions with affine connections on manifolds. Riemannian and non-Riemannian approaches was interviewed too. and we have examples in affine convexity, optimization problems.

REFERENCES

[1] P. A. ABSIL, C. G. BAKER and K. A. GALLIVAN, Trust-region methods on Riemannian manifolds, *Foundations of Computational Mathematics*, **7**, 3 (2007), pp. 303-330.

[2] P. A. ABSIL, R. MAHONY and R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, 2009.

[3] R. L. ADLER, et al., Newton's method on Riemannian manifolds and a geometric model for the human spine, *IMA Journal of Numerical Analysis*, **22**, 3 (2002), pp. 359-390.

[4] L. A. AKÉ and M. SÁNCHEZ, Compact affine manifolds with precompact holonomy are geodesically complete, *Journal of Mathematical Analysis and Applications*, **436**, 2 (2016), pp. 1369-1371.

[5] C. A. BROYDEN, A new double-rank minimisation algorithm, report, American Mathematical Society Notices, Providence, RI 02940-2213, 1969.

[6] C. G. BROYDEN, The convergence of a class of double-rank minimization algorithms 1, general considerations, *IMA Journal of Applied Mathematics*, **6**, 1 (1970), pp. 76-90.

[7] I. BRACE and J. MANTON, An improved BFGS-on-manifold algorithm for computing weighted low rank approximations, *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, 2006.

[8] N. BOUMAL and P. A. ABSIL, *RTRMC: A Riemannian trust-region method for low-rank matrix completion*, Proceedings of the 24th International Conference on Neural Information Processing Systems, (2011), pp. 406-414.

[9] N. BOUMAL, *Optimization and Estimation on Manifolds*, PhD Thesis, Universite Catholique de Louvain, Institute of information and communecation technologies, electronics and applied mathematics ICTEAM, 2014.

[10] R. W. BROCKETT, Differential geometry and the design of gradient algorithms, *Proc. of Sympo. in Pure Math*, **54** Part I (1993), pp. 69-92.

[11] I. CHAVEL, *Riemannian Geometry: A Modern Introduction*, Cambridge University Press, 2006.

[12] M. P. DO CARMO, *Riemannian Geometry*, Bikhausen, Boston, 1992.

[13] J. P. DEDIEU, P. PRIOURET and G. MALAJOVICH, Newton's method on Riemannian manifolds: covariant alpha theory, *IMA Journal of Numerical Analysis*, **23**, 3 (2003), pp. 395-419.

[14] A. EDELMAN, T. A. ARIAS and S.T. SMITH, The geometry of algorithms with orthogonality constraints, *SIAM journal on Matrix Analysis and Applications*, **20**, 2 (1998), pp. 303-353.

[15] R. FLETCHER, A nonlinear programming problem in statistics (educational testing), *SIAM Journal on Scientific and Statistical Computing*, **2**, 3 (1981), pp. 257-267.

[16] R. FLETCHER, A new approach to variable metric algorithms, *The Computer Journal*, **13**, 3 (1970), pp. 317-322.

[17] D. GOLDFARB, A family of variable-metric methods derived by variational means, *Mathematics of Computation*, **24**, 109(1970), pp. 23-26.

[18] D. GABAY, Minimizing a differentiable function over a differential manifold, *Journal of Optimization Theory and Applications*, **37**, 2(1982), pp. 177-219.

[19] U. HELMKE and J. B. MOORE, *Optimization and Dynamical Systems*, Springer Science and Business Media, 2012.

[20] U. HELMKE and J. B. MOORE, *Optimization and Dynamical Systems*, *Communications and Control Engineering Series*, Springer-Verlag London Ltd., London, 1994.

[21] D. JIANG, J. B. MOORE and H. JI, Self-concordant functions for optimization on smooth manifolds, *Journal of Global Optimization*, **38**, 3 (2007), pp. 437-457.

[22] H. JI, *Optimization Approaches on Smooth Manifolds*, The Australian National University, 2007.

[23] N. A. KARMARKAR, New polynomial-time algorithm for linear programming, *Proceedings of the sixteenth annual ACM symposium on Theory of computing, ACM*, 1984.

[24] J. M. LEE, *Riemannian Manifolds: An Introduction to Curvature*, Springer Science and Business Media, 2006.

[25] D. G. LUENBERGER, The gradient projection method along geodesics, *Management Science*, **18**, 11 (1972), pp. 620-631.

[26] D. G. LUENBERGER, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Reading, MA, 1973.

[27] R. MAHONY, *Optimization Algorithms on Homogeneous Spaces*, Australian National University, 1994.

[28] R. MAHONY, The constrained Newton method on a Lie group and the symmetric eigenvalue problem, *Linear Algebra and Its Applications*, **248** (1996), pp. 67-89.

[29] J. H. MANTON, Optimization algorithms exploiting unitary constraints, *IEEE Transactions on Signal Processing*, **50**, 3 (2002), pp. 635-650.

[30] J. H. MANTON, On the various generalisations of optimisation algorithms to manifolds, *Sixteenth International Symposium on Mathematical Theory of Networks and Systems*, Belgium, Citeseer, 2004.

[31] Y. NESTEROV and A. NEMIROVSKII, *Interior-point Polynomial Algorithms in Convex Programming*, SIAM, 1994.

[32] Y. E. NESTEROV and M. J. TODD, On the Riemannian geometry defined by self-concordant barriers and interior-point methods, *Foundations of Computational Mathematics*, **2**, 4 (2002), pp. 333-361.

[33] B. O'NEILL, *Semi-Riemannian Geometry With Applications to Relativity*, Academic press, 1983.

[34] A. S. RASHEED, C. UDRIŞTE and I. TEVY, Barrier algorithm for optimization on affine manifolds, *BSG Proceedings*, **24** (2017), pp. 47-63.

[35] W. RING and B. WIRTH, Optimization methods on Riemannian manifolds and their application to shape space, *SIAM Journal on Optimization*, **22**, 2 (2012), pp. 596-627.

[36] D. F. SHANNO, Conditioning of quasi-Newton methods for function minimization, *Mathematics of Computation*, **24**, 111 (1970), pp. 647-656.

[37] S. T. SMITH, Optimization techniques on Riemannian manifolds, *Fields Institute Communications*, **3**, 3 (1994), pp. 113-135.

[38] S. T. SMITH, *Geometric optimization methods for adaptive filtering*, Harvard University, 1993.

[39] C. UDRISTE, Optimization methods on Riemannian manifolds, *Proceedings of IRB International Workshop*, Monteroduni, Italy. 1995.

[40] C. UDRIŞTE, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Kluwer Academic Publishers, Netherlands, 1994.

[41] C. UDRIŞTE, I. TEVY, H. BONNEL and A. S. RASHEED, Bilevel disjunctive optimization on affine manifolds, Chapter in: *Poom Kumam. Optimization Alorithms* InTech Open, Croatia, Rijeka, 2018, ISBN 978-953-51-6060-1.

[42] L. VANDENBERGHE and S. BOYD, Semidefinite programming, *SIAM review*, **38**, 1(1996), pp. 49-95.

[43] M. H. WRIGHT, Interior methods for constrained optimization, *Acta Numerica*, **1** (1992), pp. 341-407.

[44] S. J. WRIGHT, Primal-dual interior-point methods, *SIAM*, 1997.