ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 6 | Issue: 12 |December 2020 || Journal DOI: 10.36713/epra2013 || SJIF Impact Factor: 7.032 ||ISI Value: 1.188

OPTIMIZATION APPROACHES ON SMOOTH MANIFOLDS

Prof.(Dr.) Shailesh Nath Pandey
Department of Mathematics
(Applied Science),
B.N. College of Engineering & Technology,
Lucknow

Shailja Dubey
Research Scholar

ABSTRACT

This paper deals with fully explore properties of the self-concordant function in Euclidean space and develop gradient-based algorithms for optimization of such function. Define the self-concordant function on Riemannian manifolds, explore its properties and devise corresponding optimization algorithms and generalize a quasi-Newton method on smooth manifolds without the Riemannian structure.

We first review the classical Riemannian approaches and then the relatively recent non-Riemannian approaches.

1- RIEMANNIAN APPROACH

1.1 -Steepest descent method on manifolds:

The steepest descent method is the simplest method for the optimization on Riemannian manifolds and it has good convergence properties but slow linear convergence rate. This method was first introduced to manifolds by Luenberger [48, 49] and Gabay [25]. In the early nineties, this method was carried out to problems in systems and control theory by Brockett [13], Helmke and Moore [34], Smith [66] and Mahony [51].

1.2 -Newton method on manifolds:

Compared against the steepest descent method, the Newton method has a faster (quadratic) local convergence rate. In 1982, Gabay extended the Newton method to a Riemannian sub-manifold of $\mathbb{R}^n$ by updating iterations along a geodesic. Other independent work has been developed to extend the Newton method on Riemannian manifolds by Smith [67] and Mahony [51, 52] restricting to the compact Lie group, and by Udriste [70] restricting to convex optimization problems on Riemannian manifolds. Edelman, Arias and Smith [19] also introduced a Newton method for the optimization on orthogonality constraints – the Stiefel and Grassmann manifolds. There is also a recent paper by Dedieu, Priouret and Malajovich [18] which studied the Newton method to find zero of a vector field on general Riemannian manifolds.

1.3-Quasi-Newton method on manifolds:

Even though the Newton's method has faster quadratic convergence rate, it requires computing the inverse of a symmetric matrix, called the Hessian consisting of the second order local information of the cost function. Therefore, it increases the computational cost. In order to avoid this problem, the quasi-Newton method in Euclidean space was presented by Davidon [17] in late 1950s. 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. Since then, various quasi-Newton methods have been introduced. However, among them, the most popular methods are the Davidon-Fletcher-Powell (DFP) [22] method andthe Broyden [15, 16] Fletcher [21] Goldfarb [28] Shanno [65] (BFGS) method.

In the early eighties, Gabay [65] firstly generalized the BFGS method to a Riemannian manifold. However, he did not give the complete proof of the convergence of his method. Recently, Brace and Manton [12] developed an improved BFGS method on the Grassmann manifold and achieved a lower computational complexity compared to Gabay's method

1.4-Conjugate gradient method on manifolds:

While considering the large scale optimization problems with sparse Hessian matrices, the quasi-Newton methods encounter difficulties. Due to avoiding computing the inverse of the Hessian, the conjugate gradient method can be used for solving such problems.

This method was originally developed by Hestenes and Stiefel [38] in the 1950s to solve large scale systems of linear equations. Then in the mid 1960s, Fletcher and Reeves [24] popularized this method to solve unconstrained optimization problems. In 1994, Smith [67] extended this method to Riemannian manifolds and later Edelman, Arias and Smith [19] applied his method specifically on the Stiefel and Grassmann manifolds.

2-SELF-CONCORDANT FUNCTIONS ON RIEMANNIAN MANIFOLDS:

Self-concordant functions play an important role in developing interior point algorithms for solving certain convex constrained optimization problems including linear programming. It is therefore natural to attempt to extend the definition of self-concordance to functions on Riemannian manifolds, and then exploit this definition to derive novel optimization algorithms on Riemannian manifolds. In fact, the self-concordant concept has been extended to Riemannian manifolds .

In that work, we can considered the convex programming problem

$$\min f_0(p) \; s.t. \; f_i(p) \leq 0 \; , \; i = 1, \ldots, m; p \in M \ldots\ldots\ldots\ldots\ldots (i)$$

Where M is a complete n-dimensional Riemannian manifold and developed a logarithmic barrier interior point method for solving it. Recall that in the Euclidean space, one approach for solving

(i) Is the barrier interior point method which uses the barrier function to enforce the constraint; this barrier function is chosen to be self-concordant. In order to extend this idea to Riemannian manifolds, it is necessary to extend the concept of self-concordant functions to Riemannian manifolds.

(ii) To this end, the concept of a self-concordant function was defined on Riemannian manifolds and some of its properties alos. Moreover, a Newton method with a step-size choice rule was proposed to keep the iterates inside the constraint and guarantee the convergence.

In this chapter, we give a precise definition of a self-concordant function on a Riemannian manifold and derive properties of self-concordant functions which will be used to develop optimization algorithms; first a damped Newton method in this chapter, then a damped conjugate gradient method. Convergence proofs of the damped Newton method are also given.

3-CONCEPTS OF RIEMANNIAN MANIFOLDS

In this section, some fundamental concepts from differential geometry are introduced. However, we do not intend to present self-contained and complete exposure, and most of the proofs are omitted.

Let an n-dimensional smooth manifold be denoted as M which is an embedded manifold in $\mathbb{R}^N$ The differential structure of M is a set of local charts covering M. Each local chart is a pair of a neighborhood and a smooth mapping from this neighborhood to an open set in Euclidean space. The tangent space of M at a point p can be denoted as $T_p M$:

It is the set of linear mappings from all smooth functions passing through the point p to real numbers, satisfying the derivative condition. For n-dimensional manifolds, the tangent space at every point is an n-dimensional vector space with origin at this point of tangency.

The normal space is the orthogonal complement of the tangent space in the ambient space. A smooth manifold M is called Riemannian manifold if it is endowed with a metric structure.

In Euclidean space, a vector can be moved parallel to itself by just moving the base of the arrow. For the manifold if a tangent vector is moved to another point on the manifold parallel to itself in its ambient space, it is generally not a tangent vector to the new point.

However, we can transport tangent vectors along paths on the manifold by infinitesimally removing the component of the transported vector in the normal space.

Assume that we want to move a tangent vector Δ along the curve $\gamma(t)$ on the manifold. Then in every infinitesimal step, we first move Δ parallel to itself in the ambient Euclidean space and then remove the normal component.

Let M denote a smooth n-dimensional geodesically complete Riemannian manifold. Recall that C^ksmooth means derivatives of the order k exist and are continuous. For convenience, by smooth, we mean C^∞that is, derivatives of all orders exist.

Let T_pM denote the tangent space at the point $p \in M$. Since M is a Riemannian manifold, it comes with an inner product $<.,.>_p$on T_pM for each $p \in M$. This induces the norm $\|.\|_p$given by

$$\| X \|_p = < X, X >_p^{\frac{1}{2}} p \text{ for } X \in T_pM.$$

There is a natural way (precisely, the Levi-Civita connection) of defining acceleration on a Riemannian manifold which is consistent with the metric structure. A curve with zero acceleration at every point is called a geodesic. Since M is geodesically complete, given a point $p \in M$ and a tangent vector $X \in T_pM$, there exists a unique geodesic $\gamma X : \mathbb{R} \to M$ such that

$$\gamma X(0) = p \text{ and } \gamma X(0) = X.$$

We therefore define an exponential map

$$\text{Exp}_p\colon T_pM \to M \text{ by}$$

$$\text{Exp}_p(X) = \gamma X(1)$$

for all $X \in T_pM$. Note that $\text{Exp}_p\, tX$ is the geodesic emanating from p in the direction X. Another consequence of M being geodesically complete is that any two points on M can be joined by a geodesic of shortest length. The distance $d(p, q)$ between two points $p, q \in M$ is defined to be the length of this minimizing geodesic.
Since the length of the curve

$$\gamma : [0, 1] \to M, \gamma(t) = Exp_p tX, is \ \| X \|_p$$

it follows that if $q = \text{Exp}_p X$ then

$$d(p, q) \leq \| X \|_p$$

where the inequality is possible if there exists a shorter geodesic connecting p and q.
If $\gamma : [0, 1] \to M$ is a smooth curve from $p = \gamma(0)$ to $q = \gamma(1)$, there is an associated linear isomorphism

$$T_{pq}\colon T_pM \to T_qM$$

called parallel transport. One of its properties is that lengths of vectors and angles between vectors are preserved, i.e.

$$\forall X, Y \in T_pM, < \gamma_{pq} X, \gamma_{pq} Y >_p < X, Y >_p.$$

For a point $p \in M$ and a tangent vector $X \in TpM$, we use $\gamma p\, \text{Exp}_p(tX)$ to denote the parallel transport from the point p to the point $\text{Exp}_p\, tX$ along the geodesic emanating from p in the direction X.

Let N be an open subset of M. Consider the function $f\colon N \to \mathbb{R}$. Given

$p \in N$ and

$X \in TpN$, the first, second and third covariant derivatives of f are defined as follows:

$$\nabla_x f(p) = \frac{d}{dt}|_{t=0} \{f(Exp_p tX)\} \dots \dots (2)$$

$$\nabla^2{}_x f(p) = \frac{d^2}{dt^2}|_{t=0} \{f(Exp_p tX)\} \dots \dots (3)$$

$$\nabla^3{}_x f(p) = \frac{d^3}{dt^3}|_{t=0} \{f(Exp_p tX)\} \dots \dots (4)$$

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 6 | Issue: 12 |December 2020 || Journal DOI: 10.36713/epra2013 || SJIF Impact Factor: 7.032 ||ISI Value: 1.188

The gradient of f at $p \in N$, denoted by $\mathrm{grad}_p f$, is defined as the unique tangent vector in T_pN such that

$$\nabla_x f(p) = \langle \mathrm{grad}_p f, X \rangle$$

for all $X \in T_pN$.

The Hessian of f at $p \in N$ is the unique symmetric bilinear form $\mathrm{Hess}_p f$ defined by the property

$$Hess_p f(X,X) = \nabla^2{}_x f(p), X \in T_pN \text{ ….. } (5)$$

Note that (5) fully defines $\mathrm{Hess}_p f$ since

$$Hess_p f(X,Y) = \frac{Hess_p f(X+Y,\ X+Y) - Hess_p f(X,X) - Hess_p f(Y,Y)}{2} \text{ ……….(6)}$$

$$X, Y \in T_pN$$

4-SELF-CONCORDANT FUNCTIONS

The definition of self-concordance to Riemannian manifolds requires carefully defining the convex set. Intuitively, the convex set on Riemannian manifolds can be determined by the geodesics connecting two points. However, there could be more than one geodesic connecting two points on Riemannian manifolds. For instance, for any two different points on the sphere, there exist two geodesics joining them. Therefore, there is no single best definition of convexity of selected subset .

The definition of convexity is concerned with all geodesics of the whole Riemannian manifolds connecting two points. On the other hand, this definition limits the definition of convex functions since in most cases, the cost functions defined on Riemannian manifolds are locally convex.

To be more general, our definition goes as follows. We say a subset N of M is convex if for any p , $q \in N$, out of all the geodesics connecting p and q, there is precisely one which is contained in N.

Then, a function $f : N \subset M \to \mathbb{R}$ is said to be convex if N is a convex set and for any geodesic $\gamma : [0,1] \to N$, the function $f o\gamma : [0,1] \to \mathbb{R}$ satisfies the usual definition of convexity, namely

$$f(\gamma\ (t) \leq (1-t) f\big(\gamma(0)\big) + t\ f\ \big(\gamma(1)\big), \qquad t \in [0,1]$$

If $f : N \to \mathbb{R}$ is C^∞ smooth and N is convex, then f is convex if and only if

$$\nabla^2{}_x f(p) \geq 0,\ 0$$

For all $p \in N$ and $X \in T_pN$.

The epigraph epi(f) of f is defined by

$$\mathrm{epi}(f) = \{ f(p\ ,\ t) \in N\, x\, \mathbb{R}\ \ / f(p \leq t \} \qquad (8)$$

A function f is said to be closed convex if its epigraph epi(f) is both convex and a closed subset of $M\ x\,\mathbb{R}$.

Definition : 4.1 Let M be a smooth n-dimensional geodesically complete Riemannian manifold.

Let $f : N \subset M \to \mathbb{R}$ be a C^3-smooth closed function. Then f is self-concordant if

1. N is an open convex subset of M;

2. f is convex on N;

3. there exists a constant M_f>0 such that the inequality

$$|\nabla^3{}_x f(p)| \le M_f\left(\nabla^2{}_x f(p)\right)^{\frac{3}{2}} \ldots\ldots\ldots\ldots (9)$$

holds for all $p \in N$ *and* $X \in T_pN$.

The reason why f is required to be closed in Definition 2 is to ensure that f behaves nicely on the boundary of N; this is shown in the following proposition.

4.2 -Proposition :

Let f: N → R be self-concordant. Let ∂(N) denote the boundary of N.Then for any $\bar{p} \in \partial(N)$ and any sequence of points $p_k \in N$ converging to $\bar{p}$ we have $f(p_k) \to \infty$.

Proof: first of all we choose for $k = 2, 3, \ldots$, define $X_k \in T_{p1}N$ to be such that $p_k := \mathrm{Exp}_{p1}X_k$ and $p_k \in N$. Since f is convex, In view of equation (7)we have

$$f\left(Exp_{p1}t\,X_k\right) \le (1 - t)\,f(p_1) + t\,f(p_k) \qquad (10)$$

where $0 \le t \le 1$.

It follows from equation (10) that if $0 < t \le 1$ then

$$f(p_1) + \frac{f\left(Exp_{p1}t\,X_k\right) - f(p_1)}{t} \le f(p_k) \ldots\ldots..(11)$$

$$As\ t \to 0\ from\ equation\ (11) we\ have$$

$$f(p_1) + \lim_{t\to 0}\frac{f\left(Exp_{p1}t\,X_k\right) - f(p_1)}{t} = f(p_1) + \nabla X_k f(p_1)$$

$$= f(p_1) + < grad_{p1}f, X_k >$$

$$\le f(p_k) \ldots\ldots\ldots\ldots..(12)$$

Therefore the sequence $\{f(p_k)\}$ is bounded below by

$$f(p_k) \ge f(p_1) + < grad_{p1}f, X_k > \cdots\ldots\ldots\ldots\ldots.(13)$$

where we recall that $X_k \in T_{p1}N$ is such that $p_k := \mathrm{Exp}_{p1}X_k$.

Assume to the contrary that the sequence $\{f(p_k), k \ge 1\}$ is bounded from above. Then it has a limit point $\bar{f}$. By considering a subsequence if necessary, we can regard it as a unique limit point of the sequence.

Let $z_k = \left(p_k, f(p_k)\right)$. Then we have

$$z_k = \left(p_k, f(p_k)\right) \to \bar{z} = \left(\bar{p}\bar{f}\right) \ldots\ldots\ldots..(14)$$

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 6 | Issue: 12 |December 2020 || Journal DOI: 10.36713/epra2013 || SJIF Impact Factor: 7.032 ||ISI Value: 1.188

By definition, z_k ∈epi(f). However, we have $\bar{z}$ ∉epi(f) since $\bar{p}$ ∉N. That is a contradiction since f is closed.

Proposition- 4.3.

Let f_i: N ⊂ M → R be self-concordant with constants M_{fi} , i = 1 , 2 and

let α , β>0. Then the function f(x) = αf_1(x) + βf_2(x) is self-concordant with the constant

$$M_f = \max\left\{\frac{1}{\sqrt{\alpha}}M_{f1}, \frac{1}{\sqrt{\beta}}M_{f2}\right\} \ldots\ldots\ldots (15)$$

Proof: let f_i; i = 1 , 2 are closed convex on N, f is closed convex on N, which can be easily write for any fixed $p \in N$ and $X \in T_pN$, we have

$$|\nabla^3{}_x f_i(p)| \le M_{fi}\left(\nabla^2{}_x f_i(p)\right)^{\frac{3}{2}}, i = 1, 2 \ldots . (16)$$

Now, consider two cases.

Case One: $\alpha\nabla^2{}_x f_1(p) + \beta\nabla^2{}_x f_2(p) = 0$

Since f_1 and f_2 are both self-concordant, we have

$$\nabla^2{}_x f_1(p) \ge 0 \ldots \ldots (17)$$

$$\nabla^2{}_x f_2(p) \ge 0 \ldots \ldots . (18)$$

Therefore from the assumption, we obtain

$$\nabla^2{}_x f_1(p) = 0 \ldots \ldots (19)$$

$$\nabla^2{}_x f_2(p) = 0 \ldots \ldots . (20)$$

By the definition of self-concordance, it follows from (19) and (20) that

$$\nabla^3{}_x f_1(p) = 0 \ldots \ldots (21)$$

$$\nabla^3{}_x f_2(p) = 0 \ldots \ldots . (22)$$

Hence it follows that

$$|\nabla^3{}_x f(p)| \le M_f\left(\nabla^2{}_x f(p)\right)^{\frac{3}{2}} \ldots \ldots . (23)$$

$$where\ M_f = \max\left\{\frac{1}{\sqrt{\alpha}}M_{f1}, \frac{1}{\sqrt{\beta}}M_{f2}\right\}$$

Case One: $\alpha\nabla^2{}_x f_1(p) + \beta\nabla^2{}_x f_2(p) \ne 0$

Denote $\omega_i = \nabla^2{}_x f_1(p)\ Since\ \omega_i > 0, i = 1, 2 \ldots . by\ the\ assumption, we\ have$

$$\frac{|\nabla^3{}_x f(p)|}{(\nabla^2{}_x f(p))^{\frac{3}{2}}} \leq \frac{|\alpha\nabla^3{}_x f_1(p)| + |\beta\nabla^3{}_x f_2(p)|}{|\alpha\nabla^2{}_x f_1(p)| + |\beta\nabla^2{}_x f_2(p)|^{\frac{3}{2}}}$$

$$\leq \frac{\alpha M_{f1}\omega_2^{\frac{3}{2}}}{[\alpha\omega_1 + \beta\omega_2]^{\frac{3}{2}}} \ldots..(24)$$

Note that the last inequality is not changing when we replace (ω_1 ,ω_2) by $(t\omega_1 , t\omega_2)$ with $t>0$. Consequently, we can assume that $\alpha\omega_1 + \beta\omega_2 = 1$. Let $\xi = \alpha\omega_1$. Then the right hand side of (24) becomes

$$\frac{M_{f1}}{\sqrt{\alpha}}\xi^{3/2} + \frac{M_{f2}}{\sqrt{\beta}}(1-\xi)^{3/2} \ldots \ldots.(25)$$

Now, consider (25) as a function in $\xi \in [0 ,1]$.
As a result, its maximum is either $\xi = 0$ or $\xi = 1$.
This completes the proof. If a function f is self-concordant with the constant M_f; then the function $M_f^{-2} f$ is self-concordant with the constant 1 as can be directly checked by a simple computation. As such, we assume $M_f = 2$ for the rest of this chapter. Such functions are called standard self-concordant.

Proposition -4.3. $\forall p \in N , Q(p ,1) \subseteq N$
This property gives a safe bound for the line search along geodesics for optimization problems so that the search will always be in the admissible domain. We need the following lemma to prove it.

Lemma 6. Let $f : N \rightarrow \mathbb{R}$ in (26) be a standard self-concordant function satisfying Assumption

2. For a point $p \in N$ and a non-zero tangent vector $X \in T_pN$, recall the definitions of $\mathrm{Exp}_p t X$ and $\tau_p \mathrm{Exp}_p(tX)$ in

Section 4.2.

Let

$$U = \{t \in \mathbb{R} \,|\mathrm{Exp}_p tX \in N\}$$

Define a function

$\emptyset : U \rightarrow \mathbb{R}$ as follows

$$\emptyset(t) := \left[\nabla^2{}_{\tau_p \mathrm{Exp}_p t(x)} x f(\mathrm{Exp}_p tX)\right]^{-\frac{1}{2}} \ldots \ldots..(26)$$

Then, the following results hold:

1. $|\emptyset'(t)| \leq 1$;
2. $If\ \emptyset(0) > 0\ then\ , (-\emptyset(0), \emptyset(0)) \subseteq U$

Proof:

It can be calculated that

$$\emptyset'(t) = -\frac{\frac{d}{dt}[\nabla^2{}_{\tau_p Exp_p t(x)^x} f(Exp_p tX)]}{2\left[\nabla^2{}_{\tau_p Exp_p t(x)^x} f(Exp_p tX)\right]^{\frac{3}{2}}}$$

$$= -\frac{[\nabla^3{}_{\tau_p Exp_p t(x)^x} f(Exp_p tX)]}{2\left[\nabla^2{}_{\tau_p Exp_p t(x)^x} f(Exp_p tX)\right]^{\frac{3}{2}}}$$

The claim 1 follows directly from the definition of self-concordant function.
we have $f(Exp_p tX)$ goes to ∞as $Exp_p tX$ approaches the boundary of *N*. It implies that the function $\nabla^2{}_{\tau_p Exp_p t(x)^x} f(Exp_p tX)$ cannot be bounded.

Therefore, we have

$$\emptyset(t) \to \infty \text{as } Exp_p tX \to \partial N \dots\dots\dots (27)$$

Since the function *f*satisfies Assumption 2, by (27), we obtain

$$U \equiv \{t|\emptyset(t) > 0\} \dots\dots\dots\dots\dots (28)$$

By the claim 1, we have

$$\emptyset(t) \geq \emptyset(0) - |t| \dots\dots (29)$$

Combining (28) and (29), it follows that

$$\left(-\emptyset(0), \emptyset(0)\right) \subseteq U \dots\dots\dots.. (30)$$

In the following, two groups of properties will be given to reveal the relationship between two different points on a geodesic. They are delicate characteristics of self-concordant functions.
In fact, they are the foundation for the polynomial complexity of self-concordant functions.

Proposition :4.4. For any p ∈ N and X_p∈ T_pN; such that for t ∈[0 ,1] the geodesic $Exp_p tX_p$ is contained in N. Let q = $Exp_p X_p$ If f : N → ℝ in equation (26) is a self-concordant function, the Following results hold:

$$\left[\nabla^2{}_{\tau pq^{X_p}} f(q)\right]^{\frac{1}{2}} \geq \frac{\left[\nabla^2{}_{X_p} f(p)\right]^{1/2}}{1 + \left[\nabla^2{}_{X_p} f(p)\right]^{1/2}} \dots\dots (31)$$

$$\nabla_{\tau pq^{X_p}} f(q) - \nabla_{X_p} f(p) \geq \frac{\nabla^2{}_{X_p} f(p)}{1 + \left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}}} \dots\dots\dots. (32)$$

$$f(q) \geq f(p) + \nabla_{X_p} f(p) + \left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}} - \ln\left(1 + \left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}}\right) \ldots\ldots..(33)$$

where τpq is the parallel transport from p to q along the geodesic $\mathrm{Exp_p tX_p}$.

Proof. Let $\emptyset(t)$ be the same function defined in Lemma 6, where one can see that $\emptyset(1) \leq \emptyset(0) + 1$.

This is equivalent to equation 31 taking into account that

$$\emptyset(0) = \left[\nabla^2{}_{X_p} f(p)\right]^{-1/2}$$

$$\emptyset(1) = \left[\nabla^2{}_{\tau pq} {}^{X_p} f(q)\right]^{-\frac{1}{2}}$$

Furthermore,

$$\nabla_{\tau pq} {}^{X_p} f(q) - \nabla_{X_p} f(p)$$

$$= \int_0^1 \nabla^2{}_{\tau p Exp_p t\, X_p} {}^{X_p} f(Exp_p t\, X_p) dt$$

$$= \int_0^1 \frac{1}{t^2} \nabla^2{}_{\tau p Exp_p t\, X_p} {}^{X_p} f\left(Exp_p t\, X_p\right) dt \ldots\ldots.(34)$$

Which leads to equation (32) using the inequality (31).
For the inequality equation (33) , notice that :

$$f(q) - f(p) - \nabla_{X_p} f(p)$$

$$= \int_0^1 \{\nabla_{\tau p Exp_p t\, X_p} {}^{X_p} f\left(Exp_p t\, X_p\right) - \nabla_{X_p} f(p)\} dt$$

$$= \int_0^1 \frac{1}{t} \left\{\nabla_{\tau p Exp_p t\, X_p} {}^{X_p} f\left(Exp_p t\, X_p\right) - \nabla_{t X_p} f(p)\right\} dt$$

$$\geq \frac{\left[\nabla^2{}_{t X_p} f(p)\right]^{\frac{1}{2}}}{t\left(1 + \left[\nabla^2{}_{t X_p} f(p)\right]^{\frac{1}{2}}\right)} dt$$

$$let\ r = \left[\nabla^2{}_{t X_p} f(p)\right]^{1/2}\ .the\ last\ integral\ becomes$$

$$\int_0^1 \frac{tr^2}{1 + tr} dt = 6 - \ln(1 + r)$$

which leads to the inequality equation (33) by replacing r with its original form.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 6 | Issue: 12 |December 2020 || Journal DOI: 10.36713/epra2013 || SJIF Impact Factor: 7.032 ||ISI Value: 1.188

Proposition :4.5. For any p ∈ N and $X_p \in W(p, 1)$, let $q = Exp_p X_p$.

If f : N → ℝ in equation (26)

is a self-concordant function, then there holds:

$$\left(1-\left[\nabla^2{}_{X_p}f(p)\right]^{\frac{1}{2}}\right)^2 \nabla^2{}_{X_p}f(p) \le \nabla^2{}_{\tau pq}{}^{X_p}f(q)$$

$$\le \frac{\nabla^2{}_{X_p}f(p)}{\left(1-\left[\nabla^2{}_{X_p}f(p)\right]^{\frac{1}{2}}\right)^2} \ldots.(35)$$

$$\nabla_{\tau pq}{}^{X_p}f(q)-\nabla_{X_p}f(p) \le \frac{\nabla^2{}_{X_p}f(p)}{\left(1-\left[\nabla^2{}_{X_p}f(p)\right]^{\frac{1}{2}}\right)^2} \ldots\ldots(36)$$

$$f(q) \le f(p)+\nabla_{X_p}f(p)-\left[\nabla^2{}_{X_p}f(p)\right]^{\frac{1}{2}}-\ln\left(1-\left[\nabla^2{}_{X_p}f(p)\right]^{\frac{1}{2}}\right) \ldots\ldots(37)$$

where τpq is the parallel transport from p to q along the geodesic $Exp_p tX_p$.

Proof:

Let $\varphi(t)$ be a function defined in the following form:

$$\varphi(t)=\frac{d^2}{dt^2}f\left(Exp_p tX_p\right) \ldots\ldots\ldots\ldots(38)$$

where $t \in [0, 1]$.

Since $Xp \in W(p, 1)$, we have $Exp_p tX_p \in N$ for all $t \in [0, 1]$.

Taking the first order derivative of φ, we obtain

$$\left|\varphi'^{(t)}\right| = \left|\frac{d^3}{dt^3}f\left(Exp_p tX_p\right)\right|$$

$$= \left|\nabla^3{}_{\tau pExp_p(t\,X_p)}{}^{X_p}f\left(Exp_p\left(t\,X_p\right)\right)\right|$$

$$\leq 2\,(\nabla^2{}_{\tau p Exp_p(tXp)} f\left(Exp_p(t\,X_p)\right)^{\frac{1}{2}} (\nabla^2{}_{\tau p Exp_p(tXp)} f\left(Exp_p(t\,X_p)\right)$$

$$= 2\left(\nabla^2{}_{\tau p Exp_p(tXp)} f\left(Exp_p(t\,X_p)\right)\right)^{\frac{1}{2}} \varphi(t)$$

$$= \frac{2}{t}\left(\nabla^2{}_{t\tau p Exp_p(tXp)} f\left(Exp_p(t\,X_p)\right)\right)^{\frac{1}{2}} \varphi(t)$$

$$\leq \frac{2}{t}\frac{t\left[\nabla^2{}_{X_p} f(p)\right]^{1/2}}{1-t\left[\nabla^2{}_{X_p} f(p)\right]^{1/2}}\varphi(t) \;\dots\dots\dots (39)$$

Integrating both sides of the inequality (39) from 0 to 1, we have

$$\left(1-\left[\nabla_{X_p} f(p)\right]^{1/2}\right)^2 \leq \frac{\varphi(1)}{\varphi(0)} \leq \frac{1}{\left(1-\left[\nabla_{X_p} f(p)\right]^{1/2}\right)^2} \dots\dots (40)$$

which is equivalent to the inequality (35).

Combining the inequality (35) and the formula (34), one obtains

$$\nabla_{\tau pq} {}^{X_p} f(q) - \nabla_{X_p} f(p) \leq \int_0^1 \frac{1}{t^2}\frac{\nabla^2{}_{tX_p} f(p)}{\left(1-\left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}}\right)^2}dt$$

$$= \frac{\nabla^2{}_{tX_p} f(p)}{1-\left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}}}$$

which proves the inequality (36).

Combining this result and using the same technique as that used in the proof of the last property, there holds:

$$f(q) - f(p) - \nabla_{\mathrm{X_p}} f(p) = \int_0^1 \nabla_{\tau p Exp_p(tXp)} X_p f\left(Exp_p t\,X_p\right)dt - \nabla_{\mathrm{X_p}} f(p)$$

$$= \int_0^1 \left\{\frac{1}{t}\left[\nabla_{\tau p Exp_p(tXp)} X_p f\left(Exp_p t\,X_p\right)dt\right] - \nabla_{\mathrm{X_p}} f(p)\right\}dt$$

$$\leq \int_0^1 \frac{\nabla^2{}_{tX_p} f(p)}{t\left(1 - \left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}}\right)} dt$$

$$= -\left[\nabla^2{}_{X_p(p)} f(p)\right]^{\frac{1}{2}} - \ln(1 - \left[\nabla^2{}_{X_p} f(p)\right]^{\frac{1}{2}})$$

As such, the inequality (37) is obtained by a simple transformation of this inequality.

REFERENCES

1. *Michael Artin, Algebra, Prentice-Hall, Upper Saddle River, New Jersey, 1991.*
2. *Garrett Birkhoff and Saunders Mac Lane, A Survey of Modern Algebra, 4th Ed.,MacMillan Publishing Co., New York, 1979. (First Ed. 1941).*
3. *Ralph L. Cohen, The Immersion Conjecture for Differentiable Manifolds, Annalsof Math., 2nd Ser., Vol. 122, No. 2 (Sep., 1985), 237{328.*
4. *C. G. Gibson, Singular Points of Smooth Mappings, Pitman Publishing, London,1979.*
5. *Martin Golubitsky and Victor Guillemin, Stable Mappings and Their Singulari-ties, Springer-Verlag, New York, 1973. (Second Printing 1980).*
6. *Victor Guillemin and Alan Pollack, Di®erential Topology, Prentice Hall, Engle-wood Cli®s, New Jersey, 1974.*
7. *Serge Lang, Di®erential Manifolds, Addison Wesley Publishing Co., Reading,Massachusetts, 1972.*
8. *Yung-Chen Lu, Singularity Theory and an Introduction to Catastrophe Theory,Springer-Verlag, New York, 1976.*
9. *John N. Mather, Stability of C1 Mappings: II. Infinitesimal Stability ImpliesStability, Annals of Math., 2nd Ser., Vol. 89, No. 2 (Mar., 1969), 254{291.7172*
10. *John Milnor, Morse Theory, Princeton University Press, Princeton, New Jersey,1963. (Second Printing 1965).*
11. *Topology from the Differential Viewpoint, University of Virgina Press,Charlottesville, Virginia, 1965. (Fourth Printing 1976).*
12. *James R. Munkres, Topology: A First Course, Prentice-Hall, Englewood Cliffs,New Jersey, 1975.*
13. *D. J. Saunders, The Geometry of Jet Bundles, London Math. Soc. Lecture Note Series 142, Cambridge University Press, Cambridge, 1989.*
14. *Michael Spivak, Calculus on Manifolds, Addison-Wesley Publishing Co., NewYork, 1965.*
15. *Shlomo Sternberg, Lectures on Di®erential Geometry, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.*
16. *Hassler Whitney, Di®erentiable Manifolds, Annals of Math., 2nd Ser., Vol. 37,No. 3 (Jul., 1963), 645{680.*
17. *The Self-Intersection of a Smooth n-Manifold in 2n-Space, Annals of Math., 2nd Ser., Vol. 45, No. 2 (Apr., 1944), 220{246.*
18. *The Singularities of a Smooth n-Manifold in 2n ¡ 1-Space, Annals of Math., 2nd Ser., Vol. 45, No. 2 (Apr., 1944), 247{293.*
19. *On Singularities of Mappings of Euclidean Spaces. I. Mappings of thePlane into the Plane, Annals of Math., 2nd Ser., Vol. 62, No. 3 (Nov., 1955).*