

Finsler notes

Yue Shen

1 Basic Background

1.1 Definition of Finsler Manifolds

Let V be a real finite dimensional vector space, and let $\{e_i\}$ be a basis for V . Furthermore, let $\frac{\partial}{\partial y^i}$ be partial differentiation in the e_i direction. Here we use *Einstein summing convention* throughout. For example, for $v \in V$, $v = v^i e_i$ [3].

Definition 1 ([3]) A function $f : V \rightarrow \mathbb{R}$ is (positively) **Homogeneous of degree $s \in \mathbb{R}$** (or **s -homogeneous**) if $f(\lambda v) = \lambda^s f(v)$ for all $v \in V$, $\lambda > 0$.

Proposition 1 ([3]) Suppose f is smooth and s -homogeneous. Then ∂f is $(s - 1)$ -homogeneous, and

$$\frac{\partial f}{\partial y^i}(v)v^i = sf(v), \quad v \in V \quad (\text{Euler's theorem}).$$

Definition 2 ([1]) A (global defined) **Finsler structure** of M is a function $F : TM \rightarrow [0, \infty)$ with the following properties:

- (i) **Regularity:** F is C^∞ on $TM \setminus \{0\}$,
- (ii) **Positive homogeneity:** $F(x, \lambda y) = \lambda F(x, y)$ for all $\lambda > 0$, i.e. F is 1-homogeneous of y .
- (iii) **strong convexity:** The $n \times n$ Hessian matrix

$$(g_{ij}) := \left(\left[\frac{1}{2} F^2 \right]_{y^i y^j} \right)$$

is positive-definite at every point of $TM \setminus \{0\}$.

The pair (M, F) is a **Finsler manifold**.

Theorem 1 ([1]) Let F be a non-negative real-valued function on \mathbb{R}^n with the properties:

- F is C^∞ on the punctured space $\mathbb{R}^n \setminus 0$.
- $F(\lambda y) = \lambda F(y)$ for all $\lambda > 0$
- The $n \times n$ matrix (g_{ij}) , where $g_{ij}(y) := [\frac{1}{2} F^2]_{y^i y^j}(y)$, is positive definite at all $y \neq 0$.

Then we have the following conclusions:

- (Positivity) $F(y) > 0$ where $y \neq 0$.
- (Triangle inequality) $F(y_1 + y_2) \leq F(y_1) + F(y_2)$, where equality holds if and only if $y_2 = \alpha y_1$ or $y_1 = \alpha y_2$ for some $\alpha \geq 0$.
- (Fundamental inequality) $w^i F_{y^i}(y) \leq F(w)$ at all $y \neq 0$, and equality holds if and only if $w = \alpha y$ for some $\alpha \geq 0$.

Remark 1

- (i) [1] The hypotheses of the above theorem define a **Minkowski norm** on \mathbf{R}^n . According to this theorem, there is no need to hypothesize that F be positive at $y \neq 0$; it is necessary so.
- (ii) [1] If the Minkowski norm satisfies $F(-y) = F(y)$, then one has the absolute homogeneity $F(\lambda y) = |\lambda|F(y)$.
- (iii) [1] In view of the first two conclusions of the above theorem, every absolutely homogeneous Minkowski norm is a norm in the sense of functional analysis.
- (iv) [3] For $u, v \in T_x M$ and $y \in T_x M \setminus \{0\}$, we have $g_y(u, v) = g_{ij}(y)u^i v^j$ where

$$g_{ij}(y) = \frac{1}{2} \frac{\partial^2 F^2}{\partial y^i \partial y^j}(y).$$

Hence g_y is bilinear.

$$(v) [6] \text{ for all } u, v \in T_x M \text{ and } y \in T_x M \setminus \{0\}, \text{ we have } g_y(u, v) = \left. \frac{1}{2} \frac{\partial^2 F^2(y+su+tv)}{\partial s \partial t} \right|_{t=s=0}$$

- (vi) [3] Suppose $u, v \in T_x M$, $y \in T_x M \setminus \{0\}$. Then

$$\begin{aligned} g_{\lambda y}(u, v) &= g_y(u, v), \quad \lambda > 0 \\ g_y(y, u) &= \left. \frac{1}{2} \frac{\partial F^2}{\partial y^i}(y)u^i \right|_{t=0} = \left. \frac{1}{2} \frac{\partial F^2(y+tu)}{\partial t} \right|_{t=0}, \\ g_y(y, y) &= F^2(y). \end{aligned}$$

- (vii) [3] $F(x, y) = 0$ if and only if $y = 0$.

1.2 Another definition of Finsler Manifolds

Definition 3 ([5]) A C_p **Banach Manifold** \mathcal{M} for $p \in \mathbb{N} \cup \{\infty\}$ is an Hausdorff topological space together with a covering by open sets $(U_i)_{i \in I}$, a family of Banach vector spaces $(E_i)_{i \in I}$ and a family of continuous mappings $(\varphi_i)_{i \in I}$ from U_i inton E_i such that

- (i) for every $i \in I$,

$$\varphi_i : U_i \longrightarrow \varphi_i(U_i) \text{ is an homeomorphism.}$$

- (ii) for every pair of indices $i \neq j$ in I ,

$$\varphi_j \circ \varphi_i^{-1} : \varphi_i(U_i \cap U_j) \subset E_i \longrightarrow \varphi_j(U_i \cap U_j) \subset E_j$$

is a C^p diffeomorphism.

Definition 4 ([5]) A Banach manifold \mathcal{V} is called **C^p -Banach Space Bundle** over another Banach manifold \mathcal{M} if there exists a Banach space E , a submersion π from \mathcal{V} to \mathcal{M} , a covering $(U_i)_{i \in I}$ of \mathcal{M} and a family of homeomorphism from $\pi^{-1}U_i$ into $U_i \times E$ such that the following diagram commutes

$$\begin{array}{ccc} \pi^{-1}U_i & \xrightarrow{\tau_i} & U_i \times E \\ \pi \searrow & & \downarrow \rho \\ & & U_i \end{array}$$

where ρ is the canonical projections from $U_i \times E$ onto U_i .

The restriction of τ on each fiber $\mathcal{V}_x := \pi^{-1}(\{x\})$, for $x \in U_i$ realizes a continuous isomorphism onto E_i . Moreover, the map $x \in U_i \cup U_j \rightarrow \tau_i \circ \tau_j^{-1}|_{\pi^{-1}(x)} \in \mathcal{L}(E, E)$ is C^p .

A submersion is a differentiable map between differentiable manifolds whose differential is everywhere surjective.

$\mathcal{L}(X, Y)$ is the space of all continuous linear maps from a topological vector X to another topological space Y .

Definition 5 ([5]) Let \mathcal{M} be a normal Banach manifold and let \mathcal{V} be a Banach Space Bundle over \mathcal{M} . A **Finsler structure** on \mathcal{V} is a continuous function $\|\cdot\|_x := \|\cdot\|_{\pi^{-1}(x)}$ is a norm on \mathcal{V}_x .

Definition 6 ([5]) Let \mathcal{M} be a normal C^p Banach manifold. $T\mathcal{M}$ equipped with a Finsler structure is called a **Finsler Manifold**.

Remark: A Finsler structure on $T\mathcal{M}$ defines in a canonical way a dual Finsler structure on $T^*\mathcal{M}$.

1.3 Relative definitions

Example 1 ([3]) Let (M, g) be a Riemannian manifold. Set $F|_{T_x M}(y) = \sqrt{g_x(y, y)}$. Then F is a Finsler norm on M .

Suppose M is a manifold. Then a **curve** is a smooth mapping $c : (a, b) \rightarrow M$ such that $(Dc)_t \neq 0$ for all t [3]. Such a curve has a **canonical lift** $\hat{c} : (a, b) \rightarrow TM \setminus \{0\}$ defined as $\hat{c}(t) = (Dc)(t)$, where Dc is the tangent of c . If (M, F) is a Finsler manifold, we define the **length** of c as

$$L(c) = \int_a^b F(\hat{c}(t))dt.$$

Definition 7 ([6]) Let U be an open subset of a Finsler manifold (M, F) . Let νU be the space of smooth vector fields on U and $\nu U^+ \subset \nu U$ be the subset of nowhere vanishing vector fields. For $V \in \nu U^+$ and for all $X, Y, Z \in \nu U$ define a trilinear form $\langle \cdot, \cdot, \cdot \rangle_V$ by

$$\langle X, Y, Z \rangle_V = \frac{1}{4} \frac{\partial^3}{\partial r \partial s \partial t} F^2(V + rX + sY + tZ)|_{r=s=t=0},$$

called the **Cartan tensor**.

The Cartan tensor is a non-Riemannian quantity. A Finsler metric reduces to a Riemannian metric if and only if its Cartan tensor vanishes [6].

Definition 8 ([6]) An **affine connection** ∇^V is a map

$$\begin{aligned}\nabla^V : \nu U \times \nu U &\rightarrow \nu U \\ (X, Y) &\rightarrow \nabla_X^V Y\end{aligned}$$

(i) linear in Y (not necessarily linear in X)

$$(ii) \quad \nabla_X^V(fY) = f\nabla_X^V Y + X(f)U$$

$$(iii) \quad \nabla_{fX}^V Y = f\nabla_X^V Y$$

for all $f \in C^\infty(U)$ and $X, Y \in \nu U$.

Definition 9 ([3]) suppose $c : (a, b) \rightarrow M$ is a curve. Then a **variation** of c is a continuous mapping $H : [a, b] \times (-\epsilon, \epsilon) \rightarrow M$ for some ϵ such that

(i) H is smooth on $(-\epsilon, \epsilon) \times (a, b)$ and with notation $c_s(\cdot) = H(\cdot, s)$

(ii) $c_0(t) = c(t)$, for all $t \in [a, b]$

(iii) $c_s(a), c_s(b) \in M$ are constants not depending on $s \in (-\epsilon, \epsilon)$.

Definition 10 ([3]) A curve c in a Finsler manifold is a **geodesic** if L is stationary at x , that is, for any variation (c_s) of c ,

$$\frac{d}{ds} L(c_s) \Big|_{s=0} = 0.$$

A Finsler manifold (M, F) is said to be **forward geodesically complete** if and only if every geodesic $\sigma : [0, 1] \rightarrow M$ parameterized to have constant Finslerian speed can be extended to a geodesic defined on $[0, \infty]$ [4].

Lemma 1 ([4]) Let (M, F) be a Finslerian manifold. At every given point $x \in M$, there exists a coordinate neighborhood U containing x and a constant $c_0 > 1$ such that

(i) for all $y \in T_x M$ and $x \in \bar{U}$, $F_x(-y) \leq c_0^2 F_x(y)$; and

(ii) for all $x_1, x_2 \in U$, $c_0^{-2} d(x_1, x_2) \leq d(x_2, x_1) \leq c_0^2 d(x_1, x_2)$.

1.4 Gradient on a Finslerian Manifold [4]

Let V be a finite-dimensional vector space and V^* the dual space. Given a Minkowski norm F on V , F is a norm in the sense that for any $y, v \in V$ and $\lambda > 0$,

$$F(\lambda y) = \lambda F(y)$$

and

$$F(y + v) \leq F(y) + F(v).$$

Define $F^* := \sup_{F(y)=1} \xi(y)$, and F^* is a Minkowski norm on V^* .

Lemma 2 ([4]) Let F be a Minkowski norm on V and F^* be the dual norm on V^* . For any vector $y \in V \setminus \{0\}$, the covector $\xi = g_y(y, \cdot) \in V^*$ satisfies

$$F(y) = F^*(\xi) = \frac{\xi(y)}{F(y)}.$$

For any covector $\xi \in V^* \setminus \{0\}$, there exists a unique vector $y \in V \setminus \{0\}$ such that $\xi = g_y(y, \cdot)$.

Given a function $f : M \rightarrow \mathbb{R}$ on a manifold M , the differential $df_x \in V^*$ at a point $x \in M$ is defined

$$df_x = \frac{\partial f}{\partial x^i}(x)dx^i,$$

and is a linear functional on $T_x M$. To connect df_x to a vector $grad f_x \in T_x M$, we need a Minkowski norm on $T_x M$. Let F be a Finsler metric on M . By definition, F_x is a Minkowski norm on $T_x M$. Assume that $df_x \neq 0$. Since the indicatrix $S := F^{-1}(1)$ is strongly convex, there is a unique unit vector $s_x \in S_x M := F_x^{-1}(1)$ and a positive number $\lambda_x > 0$ such that

$$W^{\lambda_x} = \{v : df_x(v) = \lambda_x\}$$

is tangent to $S_x M$ at s_x . By Lemma 2,

$$df_x(v) = \lambda_x g_{s_x}(s_x, v)$$

where

$$F_x^*(df_x) = df(s_x) = \lambda_x g_{s_x}(s_x, s_x) = \lambda_x.$$

Define

$$grad f_x := \lambda_x s_x = F^*(df_x)s_x.$$

Then we can write

$$df_x(v) = g_{grad f_x}(grad f_x, v), \quad v \in T_x M.$$

Definition 11 ([4]) Let $f : M \rightarrow \mathbb{R}$ be a C^1 function and consider the following set: $[\eta_1 < f < \eta_2] := \{x \in M : \eta_1 < f(x) < \eta_2\}, \infty < \eta_1 < \eta_2 < +\infty$. f is said to have the **Kurdyka-Łojasiewicz** property at $x \in \text{domain}(f)$ if there exist $\eta \in (0, \infty]$, a neighborhood U of \bar{x} and a continuous concave function $\phi : [0, \eta] \rightarrow \mathbb{R}_+$ such that:

- (i) $\phi(0) = 0$, $\phi \in C^1(0, \eta)$ and, for all $s \in (0, \eta)$, $\phi'(s) > 0$;
- (ii) for all $x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta]$, the Kurdyka-Łojasiewicz inequality holds

$$\phi'(f(x) - f(\bar{x}))F(grad f(x)) \geq 1.$$

We call f a KL function if it satisfies the Kurdyka-Łojasiewicz inequality at each point of $\text{domain}(f)$.

2 Applications of Finsler Manifolds

2.1 proximal point method on Finslerian Manifolds and the "Effort-Accuracy" Trade off[4]

Let M be a complete Finslerian manifold and let $f : M \rightarrow \mathbb{R}$. We will consider the optimization problem

$$\min f(x), \quad x \in M.$$

The proximal point algorithm in Finslerian manifold generates, for a starting point $x_0 \in M$, a sequence $\{x_k\} \subset M$ by the iteration

$$x_{k+1} = \arg \min_{z \in M} \{f(z) + \frac{1}{\lambda} C_{x_k}(z)\} \quad (1)$$

with $C_{x,k} : M \rightarrow \mathbb{R}$ defined by

$$C_{x_k}(z) = \frac{1}{2} d^2(x_k, z)$$

where d is the Finslerian distance and λ is a sequence of positive numbers.

Proposition 2 ([4]) *Let $\{x_k\}$ be the sequence concerning (1). Then (x_k) is well defined. Moreover:*

$$f(x_{k+1}) + \frac{1}{2\lambda} d^2(x_k, x_{k+1}) \leq f(x_k)$$

and consequently $\sum_{k=0}^{\infty} d^2(x_k, x_{k+1}) < \infty$.

Lemma 3 ([4]) *Let (x_k) be the sequence generated by (1) and $x_{k_0} \in B(\tilde{x}, \rho) \subset U$, where the neighborhood U is given by **Lemma 1**. Then*

$$F(\text{grad } f(x_{k_0})) \leq c_0^2 t_1^{-1} d(x_{k_0-1, x_{k_0}}),$$

where $c_0 > 1$ and $t_1 < \lambda_{k_0}$.

Lemma 4 ([4]) *Let $\{a_k\}$ be a sequence of positive numbers such that*

$$\sum_{k=1}^{+\infty} \frac{a_k^2}{a_{k-1}} < +\infty.$$

Then $\sum_{k=1}^{+\infty} a_k < +\infty$.

Lemma 5 ([4]) *Let (x_k) be the sequence generated by (1), $f : M \rightarrow \mathbb{R}$ a C^1 function, \tilde{x} a accumulation point of (x_k) and f satisfies the Kurdyka-Łojasiewica inequality at \tilde{x} . Let $a = \frac{1}{2}t_2$, $b = \frac{c_0^2}{t_1}$ constants and $\rho > 0$ such that $B(\tilde{x}, \rho) \subset U$, where U is given by **Lemma 1**. Then there exists $k_0 \in \mathbb{N}$ such that*

$$f(\tilde{x}) < f(x_k) < f(\tilde{x}) + \eta, \quad k \geq k_0,$$

$$d(\tilde{x}, x_{k_0}) + 2\sqrt{\frac{f(x_{k_0}) - f(\tilde{x})}{a}} + \frac{b}{a}\phi(f(x_{k_0}) - f(\tilde{x})) < \rho.$$

Moreover,

$$\frac{b}{a}[\phi(f(x_{k_0}) - f(\tilde{x})) - \phi(f(x_{k_0+1}) - f(\tilde{x}))] \geq \frac{d^2(x_{k_0}, x_{k_0+1})}{d(x_{k_0-1}, x_{k_0})}.$$

In particular, if $x_k \in B(\tilde{x}, \rho)$ for all $k \geq k_0$, then $\sum_{k=k_0}^{+\infty} d(x_k, x_{k+1}) < \infty$ and thus the sequence (x_k) converges to \tilde{x} .

Lemma 6 ([4]) Let (x_k) be the sequence concerning to (1) and assume that assumptions of **Lemma 5** hold. Then, there exists a $k_0 \in \mathbb{N}$ such that

$$x_k \in B(\tilde{x}, \rho), \quad k > k_0.$$

Theorem 2 ([4]) Let U , η and $\phi : [\eta] \rightarrow \mathbb{R}_+$ be the objections appearing in the **Definition 11**. Assume that $x_0 \in M$, $\tilde{x} \in M$ is an accumulation point of the sequence (x^k) , $\rho > 0$ such that $B(\tilde{x}, \rho) \subset U$ and f satisfies the Kurdyka-Łojasiewicz inequality of \tilde{x} . Then there exists $k_0 \in \mathbb{N}$ such that

$$\sum_{k=k_0}^{\infty} d(x_k, x_{k+1}) < +\infty.$$

Moreover, $f(x_k) \rightarrow f(\tilde{x})$, as $k \rightarrow +\infty$, and the sequence (x_k) converges to \tilde{x} and \tilde{x} is a critical point of f .

Theorem 3 ([4]) Assume hypotheses of **Lemma 6**. Assume further that (x_k) converges to x^* and that f has the Kurdyka-Łojasiewicz property at x^* with $\phi(s) = cs^{1-\theta}$, $\theta \in [0, 1)$, $c > 0$. Then the following estimations hold:

(i) If $\theta = 0$ then the sequence (x_k) converges in a finite number of steps;

(ii) If $\theta \in (0, \frac{1}{2}]$ then there exist $b_0 > 0$ and $\zeta \in [0, 1)$ such that

$$d(x_k, x^*) \leq b_0 \zeta^k;$$

(iii) If $\theta \in (\frac{1}{2}, 1)$ then there exist $\xi > 0$ such that

$$d(x_k, x^*) \leq \xi k^{-\frac{1-\theta}{2\theta}-1}.$$

References

- [1] David Bao, S-S Chern, and Zhongmin Shen. *An introduction to Riemann-Finsler geometry*, volume 200. Springer Science & Business Media, 2012.
- [2] G Bercu. Gradient methods on finsler manifolds. In *Proceedings of the Workshop on Global Analysis, Differential Geometry and Lie Algebras*, pages 230–233, 2000.

- [3] Matias Dahl. A brief introduction to finsler geometry, 2006.
- [4] JX Cruz Neto, P Roberto Oliveira, PA Soares, and Antoine Soubeyran. Proximal point method on finslerian manifolds and the ?effort–accuracy? trade-off. *Journal of Optimization Theory and Applications*, 162(3):873–891, 2014.
- [5] Tristan Riviere. Minmax methods in the calculus of variations of curves and surfaces. *Minicourse Columbia University*, 2016.
- [6] Katsuhiro Shiohama and Bankteshwar Tiwari. The global study of riemannian-finsler geometry. In *Geometry in History*, pages 581–621. Springer, 2019.