FINSLER METRICS INDUCED BY A SIMILARITY FUNCTION

NISACHON KUMANKAT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT
FOR THE DEGREE OF MASTER OF SCIENCE IN APPLIED MATHEMATICS
DEPARTMENT OF MATHEMATICS FACULTY OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2020

KMITL-2020-SC-M-001-036

FINSLER METRICS INDUCED BY A SIMILARITY FUNCTION

NISACHON KUMANKAT

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT
FOR THE DEGREE OF MASTER OF SCIENCE IN APPLIED MATHEMATICS
DEPARTMENT OF MATHEMATICS FACULTY OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2020

KMITL-2020-SC-M-001-036

Thesis Title	Finsler Metrics Induced by a Similarity Function
Student Name	Miss Nisachon Kumankat
Student ID	61605023
Degree	Master of Science (Applied Mathematics)
Department	Mathematics
Year	2020
Thesis Advisor	Assoc.Prof.Praiboon Pantaragphong

Abstract

In this research, we study the geometric properties of a topological space endowed with a similarity. We discuss its relation with weighted quasi-metrics and Finsler metrics of Randers type. We consider some applications to bioinformatics and computer science by relating similarities to Dynamic programming algorithms. We show that the space continuing the real world data is non-symmetric and non-linear.

Keywords : dynamic programming, Finsler metrics, similarity function, weighted quasi-metric space

Acknowledgements

Immeasurable appreciation and deepest gratitude for the advices, help and support are extended to the following persons who have contributed to making this thesis possible.

First of all, I would like to express my sincere gratitude to my advisor Associate Professor Praiboon Pantaragphong for the continuous support of my M.Sc. study and related research, for his patience, motivation, and immense knowledge. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing. Furthermore I would also like to express my sincere gratitude to Professor Dr.Sorin V. Sabau from Tokai university for introducing me to the topic as well for the useful comments, remarks and engagement through the learning process of this master thesis. I would like to express my sincere gratitude to Dr.Arnon Ploymukda for giving me advice in all the time of the research writing. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my M.Sc. study.

Besides my advisor, I would like to thank the rest of my thesis committee: Assistant Professor Dr.Sukrawan Mavecha and Associate Professor Dr.Jessada Tariboon from King Mongkut's University of Technology North Bangkok, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

I am sincerely grateful to financial supports from the Graduate Study of the Faculty of Science, King Mongkut's Institute of Technology Ladkrabang Fund.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Nisachon Kumankat

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Inception and Importance

Metric spaces, symmetric distances are used in different fields of pure mathematics like analysis, geometry and so on, as well as in applied mathematics, for instance in computer science, bioinformatics, data analysis and so on. These are the natural generalization of Euclidean and Riemannian spaces. On the other hand, non-symmetric distances, Minkowski norms, and Finsler metrics are also widely used in the analysis, differential geometry, data analysis and so on.

We claim that symmetric distances, Euclidean and Riemannian metrics are just convenient approximations (usually obtained by averaging) of the real world. The real world, based on real data measurements is highly non-symmetric and non-linear. Of course, proving such a fact in its most generality is a very complex and difficult task, beyond the purpose of this research. However, we will argue have that the similarity induced by the dynamic programming algorithm Needleman-Wunsch is actually equivalent in mature to non-symmetric distances (so-called quasi-distances) and Finsler metrics.

Indeed, our main statement in this research is that the following motions are equivalent in mature :

- Symmetric similarity function,
- Weighted quasi-distance,
- Finsler metrics of Randers type with reversible geodesics.

In [12] author showed every weighted quasi-metric space can be identified with a subspace, which is constructed from a metric space and presented a method to construct a weighted quasi-metric space, which is the graph of function determined on metric space, and showed every weighted quasi-metric space occur in this way. In [6] showed the relation between weighted quasi-metric spaces and Finsler spaces and presented the induced metric of a Randers space with reversible geodesics is a weighted quasi-metric space. In [7] author study some important geometric properties of weighted quasi-metric space and discuss the embedding of quasi-metric spaces with generalized weight.

The topic studied in the present research is very important for analysis, geometry, computer science, data analysis, bioinformatics and so on because it shows that at least partially, the reality we are living in, is non-symmetric, non-linear, non-homogeneous, and the study of data sets from the real world is actually equivalent to the use of weighted quasi-metrics on topological spaces, or of a Finsler metric of

Randers type on smooth manifolds.

Our research is organized as follows. In section 2 we introduce the similarity function on a topological space X and study its geometrical properties. Then, in section 3 we study the relation with embeddings and bundle, and finally Finsler metric. In the final section, we show the relation of the mathematical concepts introduced and studied in the previous sections with computer science and bioinformatics. Here is where we obtain our main conclusion that the world is of Finsler type.

1.2 Objectives of the Study

1) To study weighted quasi-metric and distances.
2) To study similarity function on a topological space and study its geometrical properties.
3) To study the relation with weighted quasi-metric space and Finsler metric.
4) To study the relation with Bioinformatics and computer science.

1.3 Scope of the Study

We investigate the geometrical and properties of a topologica space, concerning weighted quasi-metric and similarity function. All Finsler spaces considered in this research are Randers type.

1.4 Benefits of the Study

1) To develop a new mathematical theory of weighted quasi metric spaces.
2) To obtain the relation of similarity function with weighted quasi-metric.
3) To obtain the relation of similarity function with Finsler metrics.
4) To obtain the relation of similarity function with Bioinformatics and computer science.

1.5 Research Methodology

1) Study the geometry of metric spaces.
2) Study the similarities and distances.
3) Study the weighted quasi metric space induces a symmetric similarity function.
4) Define Finsler metric induced by a similarity function.

5) We introduce the similarity function on a topological space and study its geometrical properties.

6) Show that the relation with embeddings and bundle, and relation to Finsler metric.

7) Show that the relation with Bioinformatics and computer science.

8) Conclude that we have solved the problem that motivates this research, make suggestions for further works and write the thesis.

Table 1.1: The research schedule

Activity	Time frame (month of year)									
	2018		2019						2020	
	8-10	11-12	1-2	3-4	5-6	7-8	9-10	11-12	1-3	4-5
Step 1	⟷									
Step 2		⟷	⟷							
Step 3			⟷	⟷						
Step 4					⟷	⟷				
Step 5						⟷	⟷			
Step 6							⟷	⟷		
Step 7								⟷	⟷	
Step 8										⟷

Chapter 2

Preliminaries

The purpose of this chapter is to provide basic concepts and tools as weighted quasi-metric space, similarity function and Finsler metrics used in the research.

2.1 Weighted quasi-metric spaces

In this section, we recall the main geometrical results of weighted quasi-metrics (see [6], [12]). In special we call the attention to the bundle representation of weighted quasi-metric, a fundamental notion for the generalization following.

Definition 2.1. Let X be a non-empty set and d a real-valued function $d : X \times X \to \mathbb{R}$ that satisfies:

(i) Positiveness : $d(x,y) \geq 0$ and $d(x,x) = 0$ if and only if $x = y$,

(ii) Symmetry : $d(x,y) = d(y,x)$,

(iii) Triangle inequality : $d(x,y) \leq d(x,z) + d(z,y)$,

for any $x, y, z \in X$, then (X,d) is called a **metric space**.

Example 2.2. (i) Let $\mathbb{R}$ be the set of all real numbers and $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a function defined as

$$d(x,y) = |x - y|, \qquad \forall x, y \in \mathbb{R}. \tag{2.1}$$

Then d is a metric on $\mathbb{R}$.

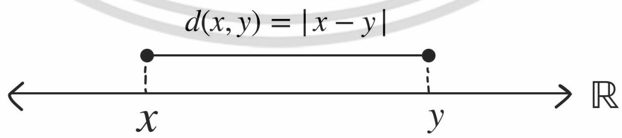

Figure 2.1: Euclidean distance in 1-dimension

(ii) Let $\mathbb{R}^n = \{(x_1, ..., x_n) : x_i \in \mathbb{R}, i = 1, 2, ..., n\}$ be the set of all n-dimensionals of real numbers and $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a function defined as

$$d(x,y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{\frac{1}{2}}, \qquad \forall x = (x_1, x_2, ..., x_n), y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n. \tag{2.2}$$

Then d is a metric on $\mathbb{R}^n$ called the **Euclidean metric** on $\mathbb{R}^n$.

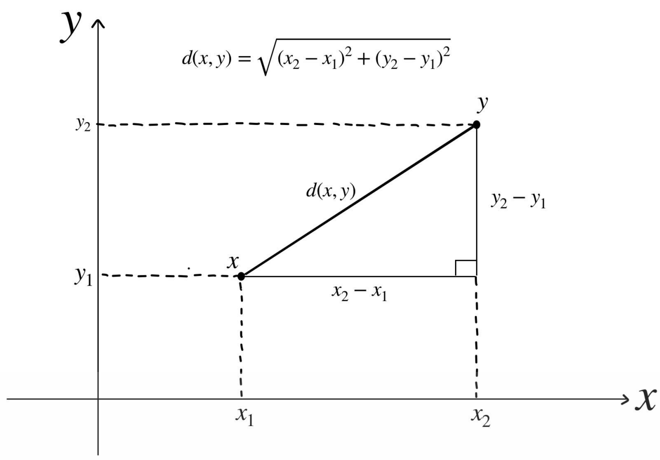

Figure 2.2: Euclidean distance in 2-dimensions

Definition 2.3. Let X be an arbitrary set and $\rho : X \times X \to \mathbb{R}$ be a real-value function satisfying the following properties for all $x, y, z \in X$:

(i) Positiveness : $\rho(x, y) > 0$ if $x \neq y$, $\rho(x, x) = 0$.

(ii) Triangle inequality : $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

(iii) Separation axiom : $\rho(x, y) = \rho(y, x) = 0 \Rightarrow x = y$.

Then (X, ρ) is called a **quasi-metric space.**

Example 2.4. (i) Let $X = [0, 1]$ and $d : X \times X \to \mathbb{R}$ be a function defined as

$$d(x,y) = \begin{cases} |x-y| & \text{, if } x, y \in X \text{ and } y \neq 0 \text{ or } x = y = 0, \\ 1 & \text{, if } y = 0 \text{ and } 0 < x \leq 1. \end{cases}$$

for all $x, y \in X$. Clearly (X, d) is a quasi-metric space because $d(0, 0.5) = |0 - 0.5| = 0.5$, while $d(0.5, 0) = 1$.

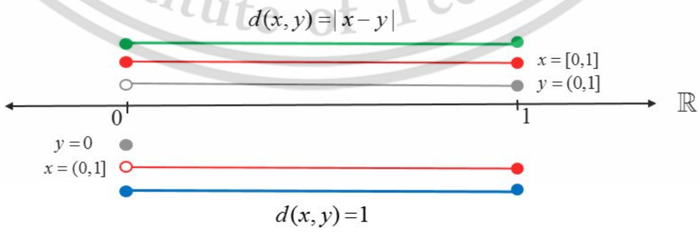

Figure 2.3: d distance in space $X = [0, 1]$

(ii) The u^L- distance in $\mathbb{R}$ is defined as

$$u^L : \mathbb{R}^{\times}\mathbb{R} \to \mathbb{R}, \qquad u^L(x, y) := max\{x - y, 0\}.$$

for all $x, y \in \mathbb{R}$. Clearly this is a quasi-metric because $u^L(1, 2) = max\{-1, 0\} = 0$, while $u^L(2, 1) = max\{1, 0\} = 1$.

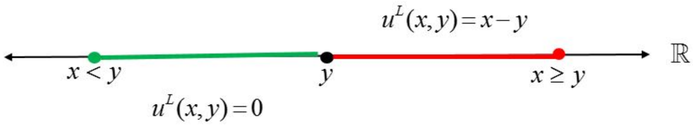

Figure 2.4: u^L-distance in 1-dimension

Moreover, an important class of quasi-metric space are the so-called **weighted quasi-metric space**.

Definition 2.5. A weighted quasi-metric space is a triple (X, ρ, w), where X is a non-empty set, $\rho : X \times X \to [0, \infty)$ and $w : X \to [0, \infty)$ satisfying the following conditions :

(i) $\rho(x, y) > 0$ if $x \neq y, \rho(x, x) = 0.$ for any $x \in X$,

(ii) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for any $x, y, z \in X$,

(iii) if $\rho(x, y) = \rho(y, x) = 0$, then $x = y$ for any $x, y \in X$,

(iv) $\rho(x, y) + w(x) = \rho(y, x) + w(y)$ for any $x, y \in X$.

The function ρ is called **quasi-metric**, and w is the **weight function**.

Example 2.6. (i) The u^L metric restricted to $[0, \infty)$, i.e. $u^L : [0, \infty) \times [0, \infty) \to \mathbb{R}$,

$$u^L(x, y) = \begin{cases} \text{y-x} & \text{, if } x \leq y \\ 0 & \text{, if } y < x \end{cases}$$

We can see that this metric u^L is a quasi-metric, with the weight function $w(x) := x$. Indeed, one can easily see that if $x \leq y$, then $u^L(x, y) \neq u^L(y, x)$ and

$$u^L(x, y) + w(x) = y - x + x = y$$
$$u^L(y, x) + w(y) = 0 + y = y.$$

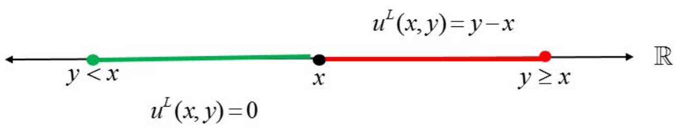

Figure 2.5: u^L-distance in 1-dimension

(ii) The metric induced by a Finsler metric of Randers type $F = \alpha + \beta$, where $\alpha(x, y) := \sqrt{a_{ij}(x)y^iy^j}, \alpha(x, y)$ Riemannian metric on $X, \beta(x, y) := b_i(x) \cdot y^i$ an exact linear 1-form on X.

Definition 2.7. If (X, ρ) is a quasi-metric space then the function $d : X \times X \to [0, \infty)$ given by

$$d(x, y) := \frac{1}{2}\left[\rho(x, y) + \rho(y, x)\right], \qquad \forall x, y \in X, \tag{2.3}$$

is called **the symmetrization** of ρ.

Lemma 2.8. If (X, ρ) is a quasi-metric space and d is symmetrization of ρ, then (X, d) is a metric space.

Proof. Let $x, y, z \in X$. We verify the conditions in definition (2.1).

(i) Positiveness:
$d(x, y) = \frac{1}{2}[\rho(x, y) + \rho(y, x)]$
Since, $\rho(x, y) \geq 0$ and $\rho(y, x) \geq 0$, it is clear that $d(x, y) \geq 0$.
Hence, $\rho(x, x) = 0$, so

$$d(x, x) = \frac{1}{2}[\rho(x, x) + \rho(x, x)] = \rho(x, x) = 0.$$

(ii) Symmetry:

$$\begin{aligned} d(x, y) &= \frac{1}{2}[\rho(x, y) + \rho(y, x)] \\ &= \frac{1}{2}[\rho(y, x) + \rho(x, y)] \\ &= d(y, x). \end{aligned}$$

(iii) The triangle inequality :

$$\begin{aligned} d(x, y) &= \frac{1}{2}[\rho(x, y) + \rho(y, x)] \\ &\leq \frac{1}{2}[\rho(x, z) + \rho(z, y) + \rho(y, z) + \rho(z, x)] \\ &= \frac{1}{2}[\rho(x, z) + \rho(z, x)] + \frac{1}{2}[\rho(z, y) + \rho(y, z)] \\ &= d(x, z) + d(z, y). \end{aligned}$$

Therefore, (X, d) is metric space. $\square$

Proposition 2.9. Using the symmetrization d of weight of quasi-metric ρ with the weight function $w : X \to [0, \infty)$ we have

$$\rho(x, y) = d(x, y) + \frac{1}{2}[w(y) - w(x)], \qquad \forall x, y \in X \tag{2.4}$$

Moreover, we have

$$\frac{1}{2}\left|w(x) - w(y)\right| \leq d(x, y), \qquad \forall x, y \in X \tag{2.5}$$

Indeed, we have

$$\begin{aligned}
\rho(x,y)+w(x) &= \rho(y,x)+w(y)\\
\rho(x,y)-\rho(y,x) &= w(y)-w(x)\\
2\rho(x,y)-(\rho(x,y)+\rho(y,x)) &= w(y)-w(x)\\
\rho(x,y) &= \frac{1}{2}[\rho(x,y)+\rho(y,x)]+\frac{1}{2}[w(y)-w(x)]\\
d(x,y) &= \rho(x,y)+\frac{1}{2}[w(y)-w(x)].
\end{aligned}$$

Proposition 2.10. If (X,ρ,w) is a weighted quasi-metric space, then the perimeter length of any geodesic triangle on X does not depend on the orientation, that is $\forall x,y,z\in X$

$$\rho(x,y)+\rho(y,z)+\rho(z,x)=\rho(x,z)+\rho(z,y)+\rho(y,x). \tag{2.6}$$

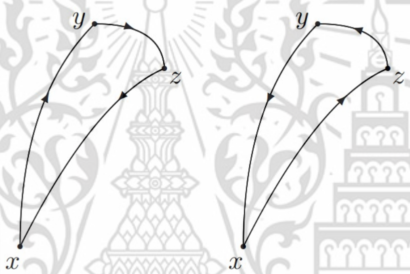

Figure 2.6: The perimeter of the triangle $\triangle xyz$ is independent of the orientation.

Proof. Let $x,y,z\in X$ and w be weight function.
From a weight ability, we have

$$\rho(x,y)+w(x)=\rho(y,x)+w(y) \tag{2.7}$$
$$\rho(y,z)+w(y)=\rho(z,y)+w(z) \tag{2.8}$$
$$\rho(z,x)+w(z)=\rho(x,z)+w(x). \tag{2.9}$$

Then (2.7)+(2.8)+(2.9), we have

$$\rho(x,y)+\rho(y,z)+\rho(z,x)=\rho(x,z)+\rho(z,y)+\rho(y,x).$$

Therefore, the proof is complete. □

Remark 2.11. It should be clear that not any quasi-metric space is weightable. In fact, it can be shown that the class of weightable quasi-metric spaces are exactly those quasi-metric spaces that satisfy relation (2.6)(see [6]).

If (X,q,w) and (Y,p,u) are two weighted quasi-metric spaces, a **morphism** of (X,q,w) into (Y,p,u) is a mapping $\varphi : X \to Y$ such that

$$p(\varphi(x),\varphi(y)) \leq q(x,y), \qquad \forall x,y \in X \tag{2.10}$$

$$u(\varphi(x)) \leq w(x), \qquad \forall x \in X. \tag{2.11}$$

We say that the morphism φ is an **isometric morphism** if in (2.10), we have equality; in this case w and $u \circ \varphi$ differ by a constant only.

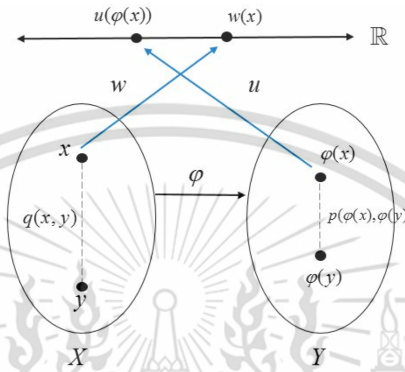

Figure 2.7: Diagram of morphism of weighted quasi-metric spaces.

Moreover, an **isomorphism** of weighted quasi-metric spaces (X,q,w) and (Y,p,u) is a bijective function $\varphi : X \to Y$ that preserves both the quasi-metric and the weight function, i.e.

$$p(\varphi(x),\varphi(y)) = q(x,y), \qquad \forall x,y \in X \tag{2.12}$$

$$u(\varphi(x)) = w(x), \qquad \forall x \in X. \tag{2.13}$$

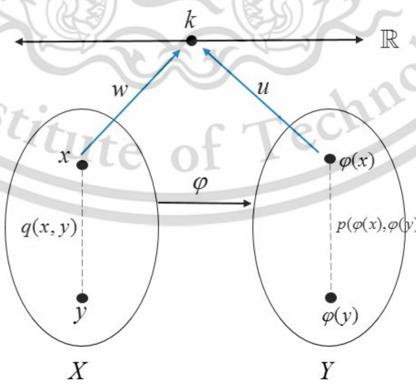

Figure 2.8: Diagram of isomorphism of weighted quasi-metric spaces.

Finally, an **embedding** of (X,q,w) into (G,Q,W) is an isomorphism of (X,q,w) onto a subspace of (G,Q,W). Here, a **subspace** (Y,p,u) **of a weighted quasi-metric space**(G,Q,W) is a subset $Y \subset G$, the function p and u are the restriction of Q and W to $Y \times Y$ and Y, respectively.

Example 2.12. (The product of a metric space with a half ray). Let us consider a metric space (X, d) and the interval $I := [0, \infty)$. Then the product space $G := X \times I$ inherits a natural structure of (generalized) weighted quasi-metric space (G, Q, W), where

$$\begin{aligned} &Q : G \times G \to I, \qquad && Q(a, b) := d(x, y) + \eta - \xi, \\ &W : G \to I, && W(a) = 2\xi, \end{aligned} \tag{2.14}$$

for any $a = (x, \xi), b = (y, \eta)$ on $G = X \times I$.

Indeed, let $a = (x, \xi), b = (y, \eta), c = (z, \zeta) \in X \times I$, then

(i) It is clear that

$$Q(a, a) = Q((x, \xi), (x, \xi)) = d(x, x) + \xi - \xi = 0.$$

(ii) Triangle inequality :

$$\begin{aligned} Q(a, b) &= Q((x, \xi), (y, \eta)) \\ &= d(x, y) + \eta - \xi \\ &\leq d(x, z) + d(z, y) + \eta - \xi \\ &= Q((x, \xi), (z, \zeta)) + \xi - \zeta + Q((z, \zeta), (y, \eta)) + \zeta - \eta + \eta - \xi \\ &= Q((x, \xi), (z, \zeta)) + Q((z, \zeta), (y, \eta)) \\ &= Q(a, c) + Q(c, b), \end{aligned}$$

So, $Q(a, b) \leq Q(a, c) + Q(c, b)$.

(iii) Suppose that $Q(a, b) = 0$ and $Q(b, a) = 0$, that is

$$\begin{aligned} Q(a, b) = Q((x, \xi), (y, \eta)) = d(x, y) + \eta - \xi \\ Q(b, a) = Q((y, \eta), (x, \xi)) = d(y, x) + \xi - \eta \end{aligned}$$

by adding these two equality we get, $2d(x, y) = 0$, whence $x = y$, so that $\xi = \eta$.

(iv) Let $W(a) = 2\xi$, we have

$$\begin{aligned} Q(a, b) + W(a) &= Q((x, \xi), (y, \eta)) + W(x, \xi) \\ &= d(x, y) + \eta - \xi + 2\xi \\ &= d(x, y) + \eta - \xi \\ &= d(y, x) + \xi - \eta + 2\eta \\ &= Q((y, \eta), (x, \xi)) + W(y, \eta) \\ &= Q(b, a) + W(b). \end{aligned}$$

Therefore, (G, Q, W) is a generalized weighted quasi-metric space.

Remark 2.13. The generalized weighted quasi-metric space (G, Q, W) constructed in Example (2.12) is sometimes called **the bundle over** (X, d)(see [6]).

Theorem 2.14. ([6]) Every weighted quasi-metric space (Y, p, w) is embeddable in a bundle over a suitable metric space (X, d).

Proof. Let metric space $(X, d) := (Y, q)$ the symmetrization of (Y, p, w).
Then, a weighted quasi-metric space (G, Q, W) the corresponding bundle, where $G := Y \times [0, \infty)$. Define a mapping $\varphi : Y \to G$ as follows:

$$\varphi(x) := \left(x, \frac{1}{2}w(x)\right), \qquad x \in X. \tag{2.15}$$

We want to show that this is an embedding.
Then, for every $x, y \in Y$, we have

$$\begin{aligned}
Q(\varphi(x), \varphi(y)) &= Q\left(\left(x, \frac{1}{2}w(x)\right), \left(y, \frac{1}{2}w(y)\right)\right) \\
&= d(x, y) + \frac{1}{2}w(y) - \frac{1}{2}w(x) \\
&= q(x, y) + \frac{1}{2}w(y) - \frac{1}{2}w(x) \\
&= p(x, y) - \frac{1}{2}w(y) + \frac{1}{2}w(x) + \frac{1}{2}w(y) - \frac{1}{2}w(x) \\
&= p(x, y).
\end{aligned}$$

Clearly we also have $\quad W(\varphi(x)) = W\left(x, \frac{1}{2}w(x)\right) = 2\left(\frac{1}{2}w(x)\right) = w(x), \quad \forall x \in Y.$
Thus, φ is an embedding. $\square$

Example 2.15. (The Graph of a function). We consider the case of the graph of a non-negative valued function $f : X \to [0, \infty)$ defined on a metric space (X, d).

Indeed, if we denote the graph of f by $G_f := \{(x, f(x)) : x \in X\}$ then (G_f, Q, W) is a naturally induced weighted quasi-metric space structure defined by

$$\begin{aligned}
&Q : G_f \times G_f \to [0, \infty), \qquad Q(a, b) := d(x, y) + f(y) - f(x), \\
&W : G_f \to [0, \infty), \qquad W(a) = 2f(x),
\end{aligned} \tag{2.16}$$

for all $a = (x, f(x)), b = (y, f(y)) \in G_f$.
Indeed, let $a = (x, f(x)), b = (y, f(y)), c = (z, f(x)) \in G_f$ we have

(i) It is clear that

$$Q(a, a) = Q((x, f(x)), (x, f(x))) = d(x, x) + f(x) - f(x) = 0.$$

(ii) Triangle inequality :

$$\begin{aligned}
Q(a,b) &= Q((x,f(x)),(y,f(y))) \\
&= d(x,y) + f(y) - f(x) \\
&\leq d(x,z) + d(z,y) + f(y) - f(x) \\
&= Q((x,f(x)),(z,f(z))) + Q((z,f(z)),(y,f(y))) \\
&= Q(a,c) + Q(c,b),
\end{aligned}$$

So, $Q(a,b) \leq Q(a,c) + Q(c,b)$.

(iii) Suppose that $Q(a,b) = 0$ and $Q(b,a) = 0$, that is

$$\begin{aligned}
Q(a,b) = Q((x,f(x)),(y,f(y))) = d(x,y) + f(y) - f(x) \\
Q(b,a) = Q((y,f(y)),(x,f(x))) = d(y,x) + f(x) - f(y)
\end{aligned}$$

by adding these two equality we get, $2d(x,y) = 0$, whence $x = y$. So that $f(x) = f(y)$.

(iv) Let $W(a) = 2f(x)$, we have

$$\begin{aligned}
Q(a,b) + W(a) &= Q((x,f(x)),(y,f(y))) + W(x,f(x)) \\
&= d(x,y) + f(y) - f(x) + 2f(x) \\
&= d(x,y) + f(y) - f(x) \\
&= d(y,x) + f(x) - f(y) + 2f(y) \\
&= Q((y,f(y)),(x,f(x))) + W(y,f(y)) \\
&= Q(b,a) + W(b).
\end{aligned}$$

Therefore, (G_f, Q, W) is a generalized weighted quasi-metric space.

Definition 2.16. ([7]) Let (X,d) be a metric space. Any function $f : X \to [0,\infty)$ is said to be a **1-Lipschitz function** if it satisfies

$$|f(x) - f(y)| \leq d(x,y), \qquad \forall x,y \in X. \tag{2.17}$$

Theorem 2.17. ([6])

1. Let (X,d) be a metric space and $f : X \to [0,\infty)$ a 1-Lipschitz function. Then the graph of f is a weighted quasi-metric space (G_f, Q, W).
2. Conversely, every weighted quasi-metric space (Y,p,w) can be constructed in this way.

Proof. For first part, by Example (2.15). Since f is a 1-Lipschitz function, we have $Q((x, f(x)), (y, f(y))) \geq 0$.
Therefore, (G_f, Q, W) is a weighted quasi-metric space with generalized weight W.
For the second part, let (Y, p, w) be a weighted quasi-metric space.
Let a metric space $(X, d) := (Y, q)$, where q is the symmetrization of p.
Define a function $f : X \to [0, \infty)$ as follows :

$$f(x) := \frac{1}{2}w(x), \qquad x \in X, \tag{2.18}$$

then f is a 1-Lipschitz function, for any $x, y \in X$, we have

$$\begin{aligned} |f(x) - f(y)| &= \frac{1}{2}|w(x) - w(y)| \\ &\leq d(x, y) \qquad\qquad \text{(by Proposition (2.9))} \end{aligned}$$

By first past, we have (G_f, Q, W) is a weighted quasi-metric space with generalized weight W, where $G_f = \{(x, f(x)) : x \in X\}$ is the graph of f and defined a mapping

$$\begin{aligned} &Q : G_f \times G_f \to [0, \infty), \quad Q((x, f(x)), (y, f(y))) := d(x, y) + f(y) - f(x), \\ &W : G_f \to [0, \infty), \qquad W((x, f(x))) = 2f(x), \end{aligned}$$

for all $x, y \in X$.
Now, by Theorem (2.14), there ia an embedding $\varphi : Y \to G$ given by

$$\varphi(x) = \left(x, \frac{1}{2}w(x)\right), \qquad x \in X, \tag{2.19}$$

where $G = Y \times [0, \infty)$ is a generalized bundle over (X, d) with a weighted quasi-metric given by

$$\begin{aligned} &Q' : G \times G \to [0, \infty), \quad Q'((x, \xi), (y, \eta)) := d(x, y) + \eta - \xi, \\ &W' : G \to [0, \infty), \qquad W'((x, \xi)) = 2\xi, \end{aligned}$$

for all $x, y \in X, \xi, \eta \in [0, \infty)$.
Since

$$\begin{aligned} \varphi(Y) &= \{\varphi(x) : x \in Y = X\} \\ &= \{(x, f(x)) : x \in X\} \\ &= G_f \end{aligned}$$

and Q, W are restrictions of Q', W' to $G_f \times G_f$ and G_f respectively, therefore, φ is an embedding of (Y, p, w) onto (G_f, Q, W), a subspace of (G, Q', W') i.e. we can identify (Y, p, w) with (G_f, Q, W). □

2.2 Finsler metrics

In this section, we will introduce Finsler spaces. Finsler spaces are viewed as regular metric spaces(see [1], [8]).

Finsler metric is a differentiable manifold M where a Minkowski function $F(x, -)$ is provided on each tangent space T_xM, that enables one to define the length of any smooth cure $\gamma : [a, b] \to M$ as

$$L(\gamma) = \int_a^b F(\gamma(t), \dot{\gamma}(t))dt.$$

Finsler metrics are more general than Remannian metrics since the tangent norms be not be induced by inner products. Every Finsler metric becomes an intrinsic quasi-metric space when the distance between two points is defined as the infimum length of the curves that join them.

The tangent bundle of a differentiable manifold M is a manifold TM which assembles all the tangent vectors in M. As a set,

$$TM = \{(x, y) | x \in M, y \in T_xM\},$$

where T_xM denotes the tangent space to M at the point x. So, an element of TM can be thought of as a pair (x, v), where x is a point in M and v is a tangent vector to M at x.

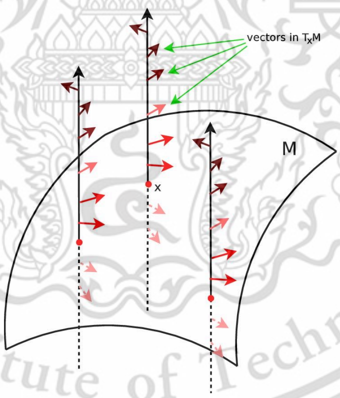

Figure 2.9: Tangent space to M at the point x

Definition 2.18. A function $F : TM \to [0, \infty)$ is called a **Finsler metric** if it has the following properties :

(i) F is C^∞ on $TM \setminus 0$,

(ii) $F(x, \lambda y) = |\lambda| F(x, y)$, for all $\lambda > 0$ and $(x, y) \in TM$,

(iii) for each $(x, y) \in TM \setminus 0$, the symmetric bilinear form g_{ij} on TM is positive definite, where

$$g_{ij} := \frac{1}{2} \frac{\partial^2}{\partial y^i \partial y^j} [F(x, y)]^2 \mid_{y^i = y^j = 0} .$$

Then the pair (M, F) is called a **Finsler space**. A Finsler metric F is said to be reversible if $F(-x, -y) = F(x, y)$ for all $(x, y) \in TM$.

Example 2.19. (Riemannian metric) Let $g = \{g_x\}_{x \in M}$, where g_x is a positive definite symmetric bilinear form in T_xM such that in local coordinates (x^i),

$$g_{ij}(x) = g_x\left(\frac{\partial}{\partial x^i}|_x, \frac{\partial}{\partial x^j}|_x\right),$$

are C^∞ functions. g is called a **Riemannian metric**. Let

$$F_x(y) = \sqrt{g_x(y, y)}, \qquad y \in T_xM. \tag{2.20}$$

Then F called a Finsler metric on M.

Example 2.20. (Rander metric) Let $\alpha(y) = \sqrt{a_{ij}(x)y^iy^j}$ and $\beta(y) = b_i(x)y^i$ be a Riemannian metric and 1-form on a manifold M. Assume that

$$\|\beta\|_x = \sup_{y \in T_xM} \frac{\beta(y)}{\alpha(y) = 1} < 1, \qquad x \in M.$$

Let a function $F : TM \to [0, \infty)$ define by

$$F(y) := \alpha(y) + \beta(y), \qquad \forall y \in TM, \tag{2.21}$$

is a Finsler metric on M. We call $F = \alpha + \beta$ a **Randers metric**.

2.3 Similarity function

In this section, we give definition and some properties about similarity function (see [10]).

Definition 2.21. Let X be a topological space. If $s : X \times X \to \mathbb{R}$ is a continous mapping such that

(i) $s(x, x) > 0$ for any $x \in X$,

(ii) $s(x, x) \geq s(x, y)$ for any $x, y \in X$,

(iii) $s(x, y) = s(x, x)$ and $s(y, x) = s(y, y)$, then $x = y$ for any $x, y \in X$,

(iv) $s(x, y) + s(y, z) \leq s(x, z) + s(y, y)$ for any $x, y, z \in X$.

Then s is called a **similarity function** on X.

The relation with quasi-metrics is well-known(see [9], [10]).

Example 2.22. Let $X = [0, \infty)$ and $s : X \times X \to \mathbb{R}$ be a function defined as

$$s(x,y) = (1+x) - \max\{x-y, 0\}, \qquad \forall x, y \in X. \tag{2.22}$$

Then s is a similarity function on X.

Indeed, let $x, y, z \in X$. We verify the conditions in the definition (2.21).

(i) It is clear that $s(x,x) = (1+x) - \max\{x-x, 0\} = 1 + x > 0$.

(ii) We will show that $s(x,x) - s(x,y) \geq 0$, we have

$$\begin{aligned} s(x,x) - s(x,y) &= (1+x) - \max\{x-x,0\} - (1+x) + \max\{x-y,0\} \\ &= \max\{x-y,0\} \\ &\geq 0 \end{aligned}$$

So, $s(x,x) - s(x,y) \geq 0$, i.e. $s(x,x) \geq s(x,y)$.

(iii) Suppose that $s(x,y) = s(x,x)$ and $s(y,x) = s(y,y)$, we have

$$\begin{gathered} s(x,y) = s(x,x) \wedge s(y,x) = s(y,y) \\ -\max\{x-y,0\} = -\max\{x-x,0\} \wedge -\max\{y-x,0\} = -\max\{y-y,0\} \\ -\max\{x-y,0\} = 0 \wedge -\max\{y-x,0\} = 0 \\ x - y \leq 0 \wedge y - x \leq 0 \end{gathered}$$

We can see that $x - y \leq 0$ and $x - y \geq 0$, then $x = y$.

(iv) We will show that $s(x,y) + s(y,z) \leq s(x,z) + s(y,y)$

Since $\quad x - z = (x-y) + (y-z) \leq \max\{x-y,0\} + \max\{y-z\}$,

and $\quad 0 \leq \max\{x-y,0\} + \max\{y-z\}$.

So, $\quad \max\{x-z,0\} \leq \max\{x-y,0\} + \max\{y-z\}$, we have

$$\begin{aligned} s(x,y) + s(y,z) &= (1+x) - \max\{x-y,0\} + (1+y) - \max\{y-z,0\} \\ &\leq (1+x) - \max\{x-z,0\} + (1+y) \\ &= (1+x) - \max\{x-z,0\} + (1+y) - \max\{y-y,0\} \\ &= s(x,z) + s(y,y) \end{aligned}$$

Therefore, s is a similarity function on X.

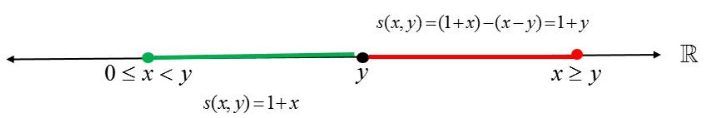

Figure 2.10: Similarity distance function in 1-dimension

Proposition 2.23. ([10]) If $s : X \times X \to \mathbb{R}$ is a similarity function on X, then $d : X \times X \to \mathbb{R}$ defined by

$$d(x, y) := s(y, y) - s(y, x), \qquad \text{for all } x, y \in X. \tag{2.23}$$

is a quasi-metric on X.

Proof. Let $x, y, z \in X$. We verify the conditions in the definition of the quasi-metric.

(i) Positiveness :
$d(x, y) = s(y, y) - s(y, x)$
Since $s(y, y) - s(y, x) \geq 0$, it is clear that $d(x, y) \geq 0$.
and $d(y, y) = s(y, y) - s(y, y) = 0$.

(ii) Triangle inequality :

$$\begin{aligned} d(x, y) &= s(y, y) - s(y, x) \\ &= s(y, y) - s(y, z) + s(y, z) - s(y, x) \\ &\leq s(y, y) - s(y, z) + s(z, z) - s(z, x) \\ &= d(z, y) + d(x, z) \\ &= d(x, z) + d(z, y) \end{aligned}$$

(iii) Separation axiom :

$$d(x, y) = 0 \wedge d(y, x) = 0 \Longrightarrow x = y$$
$$s(y, y) - s(y, x) = 0 \wedge s(x, x) - s(x, y) = 0 \Longrightarrow x = y$$
$$s(y, y) = s(y, x) \wedge s(x, x) = s(x, y) \Longrightarrow x = y$$

Therefore, (X, d) is a quasi-metric on X. □

Proposition 2.24. Let $s : X \times X \to \mathbb{R}$ is a similarity function on X. If s is a symmetric, i.e. $s(x, y) = s(y, x)$ for all $x, y \in X$, then (X, d) is a weighted quasi-metric space with weight function $w : X \to \mathbb{R}$, $w(x) = s(x, x)$.

Proof. By proposition (2.23) we have $d(x, y) = s(y, y) - s(y, x)$.
Let $x, y, z \in X$. We verify the conditions in definition of weighted quasi-metric.

(i) Positiveness :
$d(x, y) = s(y, y) - s(y, x)$
Since $s(y, y) - s(y, x) \geq 0$, it is clear that $d(x, y) \geq 0$.
and $d(y, y) = s(y, y) - s(y, y) = 0$.

(ii) Triangle inequality :

$$\begin{aligned} d(x,y) &= s(y,y) - s(y,x) \\ &= s(y,y) - s(y,z) + s(y,z) - s(y,x) \\ &\le s(y,y) - s(y,z) + s(z,z) - s(z,x) \\ &= d(z,y) + d(x,z) \\ &= d(x,z) + d(z,y) \end{aligned}$$

(iii) Separation axiom :

$$d(x,y) = 0 \wedge d(y,x) = 0 \Longrightarrow x = y$$
$$s(y,y) - s(y,x) = 0 \wedge s(x,x) - s(x,y) = 0 \Longrightarrow x = y$$
$$s(y,y) = s(y,x) \wedge s(x,x) = s(x,y) \Longrightarrow x = y$$

(iv) Let $w : X \to \mathbb{R}, w(x) = s(x,x)$, we have

$$\begin{aligned} d(x,y) + w(x) &= s(y,y) - s(y,x) + s(x,x) \\ &= s(y,y) - s(x,y) + s(x,x) \\ &= s(x,x) - s(x,y) + s(y,y) \\ &= d(y,x) + w(y) \end{aligned}$$

Therefore, (X,d) is a weighted quasi-metric space. □

We will consider in the following only symmetric similarity function.

Remark 2.25. (1) Observe that the quasi-distance d and the weight function w are determined only by the similarity function s.

(2) Let we assume the quasi-distance d is actually a distance function, i.e. $d(x,y) = d(y,x)$ for all $x,y \in X$, it follows

$$s(y,y) - s(y,x) = s(x,x) - s(x,y) \tag{2.24}$$

and if we take into account that s is symmetric, then we obtain

$$s(x,x) = s(y,y) \qquad \text{for all} \qquad x,y \in X. \tag{2.25}$$

In other words, the quasi-distance induced by a similarity function is a distance if and only if the similarity is the same on the diagonal.

(3) The symmetrized distance induced by a quasi-distance d is

$$\rho(x,y) := \frac{1}{2}[d(x,y) + d(y,x)] = \frac{1}{2}[s(x,x) + s(y,y)] - s(x,y). \tag{2.26}$$

Conversely, a weighted quasi-metric space induces a symmetric similarity function. Indeed, we have the following proposition hold

Proposition 2.26. Let (X, d) is be a weighted quasi-metric space with the weight function $w : X \to (0, \infty)$. Then the mapping $s : X \times X \to \mathbb{R}$,

$$s(x, y) := w(x) - d(y, x) \qquad \text{for all } x, y \in X, \tag{2.27}$$

is a symmetric similarity function on X.

Proof. Let $x, y, z \in X$. We verify the conditions in definition (2.21), we have

(i) Since $w : X \to \mathbb{R}$ is weight function, then $w(x) > 0$
and $d(x, x) = 0$
It is clear that $s(x, x) = w(x) - d(x, x) = w(x) > 0$

(ii) We will show that $s(x, x) - s(x, y) \geq 0$, we have

$$\begin{aligned} s(x, x) - s(x, y) &= w(x) - d(x, x) - w(x) + d(y, x) \\ &= -d(x, x) + d(y, x) \\ &= d(y, x) \end{aligned}$$

Since $d(y, x) \geq 0$
So, $s(x, x) - s(x, y) \geq 0$, i.e. $s(x, x) \geq s(x, y)$.

(iii) Suppose that $s(x, y) = s(x, x)$ and $s(y, x) = s(y, y)$, we have

$$\begin{aligned} s(x, y) = s(x, x) &\wedge s(y, x) = s(y, y) \\ w(x) - d(y, x) = w(x) - d(x, x) &\wedge w(y) - d(x, y) = w(y) - d(y, y) \\ d(x, x) - d(y, x) = w(x) - w(x) &\wedge d(y, y) - d(x, y) = w(y) - w(y) \\ d(y, x) = 0 &\wedge d(x, y) = 0 \end{aligned}$$

since $d(x, y) = 0 = d(y, x)$, thus $x = y$.

(iv) We will show that $s(x, y) + s(y, z) \leq s(x, z) + s(y, y)$, we have

$$\begin{aligned} s(x, y) + s(y, z) &= w(x) - d(y, x) + w(y) - d(z, y) \\ &= w(x) + w(y) - d(z, y) - d(y, x) \\ &\leq w(x) + w(y) - d(z, x) \\ &= w(x) - d(z, x) + w(y) - 0 \\ &= w(x) - d(z, x) + w(y) - d(y, y) \\ &= s(x, z) + s(y, y) \end{aligned}$$

Therefore, (X, s) is a symmetric similarity function on X. □

Example 2.27. Let us consider the metric space (S, ρ) and the interval $I := (0, \infty)$. It is known that the product space $G := S \times I$ interits a natrural structure of generalized weighted quasi-metric structure (G, Q, W), where

$$\begin{aligned} Q &: G \times G \to [0, \infty), & Q(a, b) &:= \rho(x, y) + \eta - \xi, \\ W &: G \to [0, \infty), & W(a) &= 2\xi, \end{aligned} \tag{2.28}$$

for any $a = (x, \xi), b = (y, \eta)$ on $G = X \times I$.
The similarity function $\varphi : G \times G \to \mathbb{R}$ induced by the weighted quasi-metric structure (G, Q, W) is given by

$$\varphi(a, b) := -\rho(x, y) + \xi + \eta, \qquad \forall a = (x, \xi), b = (y, \eta) \in G. \tag{2.29}$$

Clearly this is a symmetric similarity function on G.
Indeed, let $a = (x, \xi), b = (y, \eta), c = (z, \zeta) \in G$, then

(i) It is clear that

$$\varphi(a, a) = \varphi((x, \xi), (x, \xi)) = -\rho(x, x) + \xi + \xi = 2\xi > 0.$$

(ii) We will show that $\varphi(a, a) - \varphi(a, b) \geq 0$, we have

$$\begin{aligned} \varphi(a, a) - \varphi(a, b) &= -\rho(x, x) + \xi + \xi + \rho(x, y) - \xi - \eta \\ &= -\rho(x, x) + \rho(x, y) + \xi - \eta \\ &= \rho(x, y) + \xi - \eta \\ &= \rho(y, x) + \xi - \eta \\ &= Q((y, \eta), (x, \xi)) \\ &= Q(b, a) \\ &\geq 0 \end{aligned}$$

So, $\varphi(a, a) - \varphi(a, b) \geq 0$.

(iii) Suppose that $\varphi(a, b) = \varphi(a, a)$ and $\varphi(b, a) = \varphi(b, b)$, that is

$$\varphi((x, \xi), (y, \eta)) = \varphi((x, \xi), (x, \xi)) \wedge \varphi((y, \eta), (x, \xi)) = \varphi((y, \eta), (y, \eta))$$
$$-\rho(x, y) + \xi + \eta = -\rho(x, x) + \xi + \xi \wedge -\rho(y, x) + \eta + \xi = -\rho(y, y) + \eta + \eta$$
$$\rho(x, x) - \rho(x, y) + \eta - \xi = 0 \wedge \rho(y, y) - \rho(y, x) + \xi - \eta = 0$$

by subtract these two equality we get, $\rho(x, x) - 2\xi = \rho(y, y) - 2\eta$, whence $\xi = \eta$, so that $x = y$. Hence $a = b$.

(iv) We will show that $\varphi(a,b)+\varphi(b,c)\leq\varphi(a,c)+\varphi(b,b)$, we have

$$\begin{aligned}\varphi(a,b)+\varphi(b,c)&=\varphi((x,\xi),(y,\eta))+\varphi((y,\eta),(z,\zeta))\\&=-\rho(x,y)+\xi+\eta-\rho(y,z)+\eta+\zeta\\&=-\rho(x,y)-\rho(y,z)+\xi+\zeta+\eta+\eta\\&\leq-\rho(x,z)+\xi+\zeta+\eta+\eta\\&=-\rho(x,z)+\xi+\zeta-\rho(y,y)+\eta+\eta\\&=\varphi((x,\xi),(z,\zeta))+\varphi((y,\eta),(y,\eta))\\&=\varphi(a,c)+\varphi(b,b)\end{aligned}$$

And we will show that φ is a symmetric, i.e. $\varphi(a,b)=\varphi(b,a)$.

$$\begin{aligned}\varphi(a,b)&=\varphi((x,\xi),(y,\eta))\\&=-\rho(x,y)+\xi+\eta\\&=-\rho(y,x)+\eta+\xi\\&=\varphi((y,\eta),(x,\xi))\\&=\varphi(b,a)\end{aligned}$$

Therefore, φ is a symmetric similarity function on G.

Example 2.28. Let (S,ρ) be a metric space and $f:S\to(0,\infty)$ is a Lipschitz function with respect to ρ. Then it is known that the graph of f, i.e. $G_f=\{(x,f(x)):x\in X\}$ has a weighted quasi-metric space structure (G_f,Q,W) given by

$$\begin{aligned}&Q:G_f\times G_f\to[0,\infty), && Q(a,b):=\rho(x,y)+f(y)-f(x),\\&W:G_f\to[0,\infty), && W(a)=2f(x),\end{aligned}\tag{2.30}$$

for any $a=(x,f(x)),b=(y,f(y))$ on G_f.

It follows that the function $\varphi_f:G_f\times G_f\to\mathbb{R}$ given by

$$\varphi_f(a,b):=-\rho(x,y)+f(x)+f(y),\tag{2.31}$$

for any $a=(x,f(x)),b=(y,f(y))\in G_f$, is a symmetric similarity function on G_f. Indeed, let $a=(x,f(x)),b=(y,f(y)),c=(z,f(z))\in G_f$, then

(i) It is clear that

$$\varphi_f(a,a)=\varphi_f((x,f(x)),(x,f(x)))=-\rho(x,x)+f(x)+f(x)=2f(x)>0.$$

(ii) We will show that $\varphi_f(a,a) - \varphi_f(a,b) \geq 0$, we have

$$\begin{aligned}
\varphi_f(a,a) - \varphi_f(a,b) &= -\rho(x,x) + f(x) + f(x) + \rho(x,y) - f(x) - f(y) \\
&= -\rho(x,x) + \rho(x,y) + f(x) - f(y) \\
&= \rho(x,y) + f(x) - f(y) \\
&= \rho(y,x) + f(x) - f(y) \\
&= Q((y,f(y)),(x,f(x))) \\
&= Q(b,a) \\
&\geq 0
\end{aligned}$$

So, $\varphi_f(a,a) - \varphi_f(a,b) \geq 0$.

(iii) Suppose that $\varphi_f(a,b) = \varphi_f(a,a)$ and $\varphi_f(b,a) = \varphi_f(b,b)$, that is

$$\varphi_f(a,b) = \varphi_f(a,a) \wedge \varphi_f(b,a) = \varphi_f(b,b)$$
$$-\rho(x,y) + f(x) + f(y) = f(x) + f(x) \wedge -\rho(y,x) + f(y) + f(x) = f(y) + f(y)$$
$$\rho(x,x) - \rho(x,y) + f(y) - f(x) = 0 \wedge \rho(y,y) - \rho(y,x) + f(x) - f(y) = 0$$

by subtract these two equality we get, $\rho(x,x) - 2f(x) = \rho(y,y) - 2f(y)$, whence $f(x) = f(y)$, so that $x = y$. Hence $a = b$.

(iv) We will show that $\varphi_f(a,b) + \varphi_f(b,c) \leq \varphi_f(a,c) + \varphi_f(b,b)$, we have

$$\begin{aligned}
\varphi_f(a,b) + \varphi_f(b,c) &= \varphi_f((x,f(x)),(y,f(y))) + \varphi_f((y,f(y)),(z,f(z))) \\
&= -\rho(x,y) + f9x) + f(y) - \rho(y,z) + f(y) + f(z) \\
&= -\rho(x,y) - \rho(y,z) + f(x) + f(z) + f(y) + f(y) \\
&\leq -\rho(x,z) + f(x) + f(z) + f(y) + f(y) \\
&= -\rho(x,z) + f(x) + f(z) - \rho(y,y) + f(y) + f(y) \\
&= \varphi_f((x,f(x)),(z,f(z))) + \varphi_f((y,f(y)),(y,f(y))) \\
&= \varphi_f(a,c) + \varphi_f(b,b)
\end{aligned}$$

And we will show that φ_f is a symmetric, i.e. $\varphi_f(a,b) = \varphi_f(b,a)$.

$$\begin{aligned}
\varphi_f(a,b) &= \varphi_f((x,f(x)),(y,f(y))) \\
&= -\rho(x,y) + f(x) + f(y) \\
&= -\rho(y,x) + f(y) + f(x) \\
&= \varphi_f((y,f(y)),(x,f(x))) \\
&= \varphi_f(b,a)
\end{aligned}$$

Therefore, φ_f is a symmetric similarity function on G_f.

We can conclude that a metric space (X, ρ) with a Lipschitz function $f : X \to \mathbb{R}$ induces a similarity function on X.

The similarity space (G_f, φ_f) constructed here is called the bundle over the metric space (S, ρ).

Chapter 3

Embedding and relation to Finsler spaces

In this chapter, we study the relation the similarity function with embeddings and bundle, and with Finsler space, we study their fundamenlai properties and show that the relation with weighted quasi-metric. All the content of this section is new and makes the core of the research.

Let (X, q, w) and (Y, p, u) are two weighted quasi-metric spaces, the function $\varphi : X \to Y$ with the properties

$$p(\varphi(x), \varphi(y)) = q(x, y), \qquad \forall x, y \in X \tag{3.1}$$

$$u(\varphi(x)) = w(x), \qquad \forall x \in X. \tag{3.2}$$

is called **embeddings** of weighted quasi-metric.

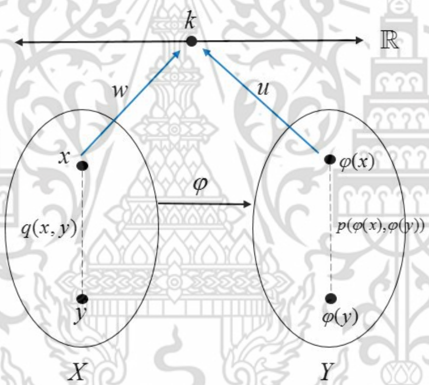

Figure 3.1: Diagram of embedding of weighted quasi-metric spaces.

Let (X, σ) and (Y, τ) be two topological spaces with similarities functions σ and τ, respectively. A continuous mapping $\varphi : X \to Y$ is called a **similarity embedding** if

$$\tau(\varphi(x), \varphi(y)) = \sigma(x, y), \tag{3.3}$$

for all $x, y \in X$.

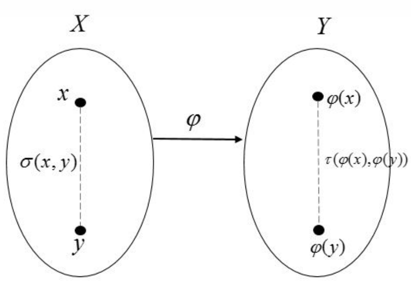

Figure 3.2: Diagram of similarity embedding.

Proposition 3.1. Let (X, q, w) and (Y, p, u) be two weighted quasi metric spaces with the associated similarities σ and τ, respectively. The continuous function $\varphi : X \to Y$ is a similarity embedding if and only if it is an embedding of weighted quasi spaces.

Proof. We assume that $\varphi : (X, \sigma) \to (Y, \tau)$ is a similarity embedding, i.e.

$$\tau(\varphi(x), \varphi(y)) = \sigma(x, y), \qquad \forall x, y \in X.$$

The weighted quasi-metric (d, w) associated to the similarity function σ on X is given by

$$d(x, y) = \sigma(y, y) - \sigma(y, x), \qquad \forall x, y \in X.$$
$$w(x) = \sigma(x, x), \qquad \forall x \in X.$$

The weighted quasi-metric $(\hat{d}, \hat{w})$ associated to the similarity function τ on Y is given by

$$\hat{d}(x, y) = \tau(y, y) - \tau(y, x), \qquad \forall x, y \in Y.$$
$$\hat{w}(x) = \tau(x, x), \qquad \forall x \in Y.$$

We compute

$$\begin{aligned}\hat{d}(\varphi(x), \varphi(y)) &= \tau(\varphi(y), \varphi(y)) - \tau(\varphi(y), \varphi(x)) \\ &= \sigma(y, y) - \sigma(y, x) \\ &= d(x, y),\end{aligned}$$

for all $x, y \in X$. Likewise,

$$\hat{w}(\varphi(x)) = \tau(\varphi(x), \varphi(x)) = \sigma(x, x) = w(x),$$

and hence it results that $\varphi : (X, d, w) \to (Y, \hat{d}, \hat{w})$ is an embedding of weighted quasi-metrics.

Conversely, we assume that $\varphi : (X, d, w) \to (Y, \hat{d}, \hat{w})$ is an embedding of weighted quasi spaces, i.e.

$$\hat{d}(\varphi(x), \varphi(y)) = d(x, y),$$
$$\hat{w}(\varphi(x)) = w(x),$$

for all $x, y \in X$. Using now relations

$$\sigma(x, y) = w(x) - d(y, x), \qquad \forall x, y \in X.$$
$$\tau(x, y) = \hat{w}(x) - \hat{d}(y, x), \qquad \forall x, y \in Y.$$

We have

$$\begin{aligned}\tau(\varphi(x), \varphi(y)) &= \hat{w}(\varphi(x)) - \hat{d}(\varphi(y), \varphi(x)) \\ &= w(x) - d(x, y) \\ &= \sigma(x, y)\end{aligned}$$

for all $x, y \in X$,therefore $\varphi : (X, \sigma) \to (Y, \tau)$ is a similarity embedding. □

Theorem 3.2. Every space with a symmetry function (X, s) is embeddable in a bundle over a suitable metric space (S, ρ).

Proof. Let a metric space (X, ρ) and let (X, s) be a space with a symmetry function s. By, Proposition (2.23) and (2.24) we have (d, w) is weighted quasi-metric with the similarity function s given by

$$d(x, y) = s(y, y) - s(y, x), \qquad \forall x, y \in X,$$
$$w(x) = s(x, x), \qquad \forall x \in X,$$

and the symmetrized associated distance s given by

$$\rho(x, y) = \frac{1}{2}[s(x, x) + s(y, y)] - s(x, y), \qquad \forall x, y \in X.$$

Then, similarity space (G_f, φ_f) the corresponding bundle, where $G_f := \{(x, f(x)) : x \in X\}$ and $f(x) := \frac{1}{2}w(x)$.
Define a mapping $\varphi : (X, s) \to (G_f, \varphi_f)$ as follows:

$$\varphi(x) := \left(x, \frac{1}{2}w(x)\right).$$

We want to show that this is an embedding.
Then, for any $x, y \in X$, we have

$$\begin{aligned}
\varphi_f(\varphi(x), \varphi(y)) &= \varphi_f\left(\left(x, \frac{1}{2}w(x)\right), \left(y, \frac{1}{2}w(y)\right)\right) \\
&= -\rho(x, y) + \frac{1}{2}w(x) + \frac{1}{2}w(y) \\
&= -\frac{1}{2}[s(x, x) + s(y, y)] + s(x, y) + \frac{1}{2}s(x, x) + \frac{1}{2}s(y, y) \\
&= s(x, y)
\end{aligned}$$

Therefore, φ is an embedding. □

Theorem 3.3. 1 Let (S, ρ) be a metric space and $f : S \to (0, \infty)$ a Lipschitz function on this metric space. Then the graph of f admits a similarity function $\varphi : G_f \times G_f \to \mathbb{R}$ that depends on ρ and f only.

2 Conversely, every similarity space (X, s) can be constructed in this way.

Proof. For statement (1), let (S, ρ) be a metric space and $f : S \to (0, \infty)$ a Lipschitz function on this metric space.
Let (G_f, Q, W) is a weighted quasi-metric space, where $G_f = \{(x, f(x)) : x \in S\}$ is the graph of f and defined a mapping

$$\begin{aligned}
Q &: G_f \times G_f \to [0, \infty), & Q(a, b) &:= \rho(x, y) + f(y) - f(x), \\
W &: G_f \to [0, \infty), & W(a) &:= 2f(x),
\end{aligned} \tag{3.4}$$

for any $a = (x, f(x)), b = (y, f(y))$ on G_f.

It follows that the function $\varphi_f : G_f \times G_f \to \mathbb{R}$ given by

$$\varphi_f(a, b) := -\rho(x, y) + f(x) + f(y), \tag{3.5}$$

for any $a = (x, f(x)), b = (y, f(y)) \in G_f$.

Now, we will show that φ_f is a symmetric similarity function on G_f.

Indeed, let $a = (x, f(x)), b = (y, f(y)), c = (z, f(z)) \in G_f$, then

(i) It is clear that

$$\varphi_f(a, a) = \varphi_f((x, f(x)), (x, f(x))) = -\rho(x, x) + f(x) + f(x) = 2f(x) > 0.$$

(ii) We will show that $\varphi_f(a, a) - \varphi_f(a, b) \geq 0$, we have

$$\begin{aligned} \varphi_f(a, a) - \varphi_f(a, b) &= -\rho(x, x) + f(x) + f(x) + \rho(x, y) - f(x) - f(y) \\ &= -\rho(x, x) + \rho(x, y) + f(x) - f(y) \\ &= \rho(x, y) + f(x) - f(y) \\ &= \rho(y, x) + f(x) - f(y) \\ &= Q((y, f(y)), (x, f(x))) \\ &= Q(b, a) \\ &\geq 0 \end{aligned}$$

So, $\varphi_f(a, a) - \varphi_f(a, b) \geq 0$.

(iii) Suppose that $\varphi_f(a, b) = \varphi_f(a, a)$ and $\varphi_f(b, a) = \varphi_f(b, b)$, that is

$$\begin{gathered} \varphi_f(a, b) = \varphi_f(a, a) \wedge \varphi_f(b, a) = \varphi_f(b, b) \\ -\rho(x, y) + f(x) + f(y) = f(x) + f(x) \wedge -\rho(y, x) + f(y) + f(x) = f(y) + f(y) \\ \rho(x, x) - \rho(x, y) + f(y) - f(x) = 0 \wedge \rho(y, y) - \rho(y, x) + f(x) - f(y) = 0 \\ -\rho(x, y) + f(y) - f(x) = 0 \wedge -\rho(y, x) + f(x) - f(y) = 0 \end{gathered}$$

by subtract these two equality we get, $\rho(x, y) + 2f(x) = \rho(y, x) + 2f(y)$, whence $f(x) = f(y)$, so that $x = y$. Hence $a = b$.

(iv) We will show that $\varphi_f(a, b) + \varphi_f(b, c) \leq \varphi_f(a, c) + \varphi_f(b, b)$, we have

$$\begin{aligned} \varphi_f(a, b) + \varphi_f(b, c) &= \varphi_f((x, f(x)), (y, f(y))) + \varphi_f((y, f(y)), (z, f(z))) \\ &= -\rho(x, y) + f9x) + f(y) - \rho(y, z) + f(y) + f(z) \\ &= -\rho(x, y) - \rho(y, z) + f(x) + f(z) + f(y) + f(y) \\ &\leq -\rho(x, z) + f(x) + f(z) + f(y) + f(y) \\ &= -\rho(x, z) + f(x) + f(z) - \rho(y, y) + f(y) + f(y) \\ &= \varphi_f((x, f(x)), (z, f(z))) + \varphi_f((y, f(y)), (y, f(y))) \\ &= \varphi_f(a, c) + \varphi_f(b, b) \end{aligned}$$

So, $\varphi_f(a, b) + \varphi_f(b, c) \leq \varphi_f(a, c) + \varphi_f(b, b)$.

And we will show that φ_f is a symmetric, i.e. $\varphi_f(a,b) = \varphi_f(b,a)$.

$$\begin{aligned}
\varphi_f(a,b) &= \varphi_f((x,f(x)),(y,f(y))) \\
&= -\rho(x,y) + f(x) + f(y) \\
&= -\rho(y,x) + f(y) + f(x) \\
&= \varphi_f((y,f(y)),(x,f(x))) \\
&= \varphi_f(b,a)
\end{aligned}$$

Therefore, φ_f is a symmetric similarity function on G_f.

For statement (2), let (X,s) be a similarity space.
Let (X,d,w) is a weighted quasi-metric space with the associated similarity s, where d and w are given by

$$\begin{aligned}
&d : X \times X \to \mathbb{R}, && d(x,y) = s(x,x) - s(y,x), \\
&w : X \to \mathbb{R}, && w(x) = s(x,x),
\end{aligned} \tag{3.6}$$

for all $x,y \in X$, and the symmetrized associated distance s given by

$$\rho(x,y) = \frac{1}{2}\left[s(x,x) + s(y,y)\right] - s(x,y), \qquad \forall x,y \in X.$$

Define a function $f : X \to (0,\infty)$ as follows :

$$f(x) := \frac{1}{2}w(x), \qquad x \in X,$$

then f is a Lipschitz function, for any $x,y \in X$, we have

$$\begin{aligned}
|f(x) - f(y)| &= \frac{1}{2}|w(x) - w(y)| \\
&\leq d(x,y) && \text{(by Proposition (2.9))}
\end{aligned}$$

By statement (1), we have (G_f,φ_f) is a similarity space.
Define a mapping $\varphi : (X,s) \to (G_f,\varphi_f)$ as follows :

$$\varphi(x) := \left(x, \frac{1}{2}w(x)\right), \qquad \forall x \in X.$$

We want to show that this is an embedding.
Then, for any $x,y \in X$, we have

$$\begin{aligned}
\varphi_f(\varphi(x),\varphi(y)) &= \varphi_f\left(\left(x,\frac{1}{2}w(x)\right),\left(y,\frac{1}{2}w(y)\right)\right) \\
&= -\rho(x,y) + \frac{1}{2}w(x) + \frac{1}{2}w(y) \\
&= -\frac{1}{2}[s(x,x)+s(y,y)] + s(x,y) + \frac{1}{2}s(x,x) + \frac{1}{2}s(y,y) \\
&= s(x,y)
\end{aligned}$$

Therefore, φ is an embedding of (X,s) onto (G_f,φ_f). $\square$

Lemma 3.4. Let M be a compact smooth manifold. If (M, ρ) is an upper and lower curvature bounded metric space, then there exists a Riemannian metric g on M whose distance function coincides with ρ.

Proof. The proof is quite obvious. Every an upper and lower curvature bounded metric space (M, ρ) constructed on an n-dimensional compact smooth manifold M can be embedded isometrically in the Euclidean space $\mathbb{R}^{2n+1}$ with the canonical metric. On the other hand, the manifold M as a submanifold in $\mathbb{R}^{2n+1}$ inherits a canonical Riemannian metric from the embedding in the Euclidean space whose distance function obviously coincides with ρ. □

Theorem 3.5. 1. If $(M, F = \alpha + \beta)$ is a simply connected Randers defined by a Riemannian metric $\alpha = \sqrt{a_{ij}(x)y^i y^j}$ and a closed 1-form β, then M is endowed with a naturally induced similarity function.

2. Conversely, let s be a symmetric, similarity function defined on a compact differentible manifold M whose associated distance ρ is an upper and lower curvature bounded. If s is differentiable, then there exists a naturally constructed Randers metric on M that depends on s only.

Proof. For statement 1, let $(M, F = \alpha + \beta)$ is a simply connected Randers defined by a Riemannian metric $\alpha = \sqrt{a_{ij}(x)y^i y^j}$ and a closed 1-form β.
Let $\gamma_{xy} \in \Gamma_{xy} := \{\gamma : [a, b] \to M | \gamma(\text{piecewise})C^{\infty} - \text{curve}, \gamma(a) = x, \gamma(b) = y\}$ be an F-geodesic and define the induced metric of a Finsler space (M, F) given by

$$d_F : M \times M \to [0, \infty), \qquad d_F(x, y) = \inf_{\gamma \in \Gamma_{xy}} \int_a^b F(\gamma(t), \dot{\gamma}(t)). \tag{3.7}$$

We have

$$\begin{aligned} d_F(x, y) &= \int_a^b \alpha(\gamma_{xy}(t), \dot{\gamma}_{xy}(t))dt + \int_a^b b_i(\gamma_{xy}(t))\dot{\gamma}^i_{xy}(t)dt \\ &= d_\alpha(x, y) + \int_{\gamma_{xy}} \beta. \end{aligned}$$

Since β is a closed 1-form on M, i.e. $d\beta = 0$ there exits a smooth function $w : M \to \mathbb{R}$ such that $\beta = dw$, we have

$$\int_{\gamma_{xy}} \beta = w\gamma(b) - w\gamma(a) = w(y) - w(x).$$

Now, it follows that M become a weighted quasi-metric space $(M, d_F, 2w)$, where

$$d_F(x, y) = d_\alpha(x, y) + w(y) - w(x), \qquad \forall x, y \in M.$$

By proposition (2.26) we have

$$\begin{aligned} s(x, y) &= 2w(x) - d_F(y, x) \\ &= 2w(x) - d_\alpha(y, x) - w(x) + w(y) \\ &= -d_\alpha(y, x) + w(x) + w(y). \end{aligned}$$

By example (2.28) we have $s(x, y) = -d_\alpha(y, x) + w(x) + w(y)$ is a symmetry similarity function.
Therefore, M is endowed with a naturally induced similarity function.

For statement 2, let s be a symmetric, similarity function defined on a compact differentible manifold M. We assume that s is a differentiable.
Define a mapping $d : M \times M \to \mathbb{R}$ as follows:

$$d(x, y) = s(y, y) - s(y, x), \qquad \forall x, y \in M,$$

is a weighted quasi-metric where a weight function $w : M \to \mathbb{R}$ given by

$$w(x) = s(x, x), \qquad \forall x \in M,$$

Then, (M, d, w) is a weighted quasi-metric induces a similarity function s.
By Lemma (3.4) we can see that there exits a Riemannian metric α on M whose distance function is exactly ρ, where ρ is a symmetrized distance induced by d, i.e.

$$\rho(x, y) = \frac{1}{2}\left[d(x, y) + d(y, x)\right] = \frac{1}{2}\left[s(x, x) + s(y, y)\right] - s(x, y), \qquad \forall x, y \in M.$$

Since s is a differentiable we can define $\beta := dw$ where w is the weight induced by s.
Therefore, $(M, F = \alpha + \beta)$ is a Randers space. $\square$

Chapter 4

Relation with Bioinformatics and Computer Science

In this chapter, we study the relation of the similarity function with bioinformatics and computer science and show that a Randers type metric whose associated distance function coincides with the weighted quasi-metric induced by a similarity function.

In order to assess the application of this theory, we start by recalling that Dynamic Programming is, at the same time, a mathematical optimization method as well as an algorithmic method in computer science. Dynamic Programming originates in the research of R. Bellman in the 1950s and it was applied eventually in many fields of science like engineering, economics, and others. In the majority cases, this method works by simplifying a much more complicated problem by dividing it into much easier small problems using a recursive way. It is known that if a problem in computer science can be solved optimally by dividing it into smaller problems and then recursively determine the optimal solutions to these small problems, then the original problem has an optimal substructure. The algorithms involving Dynamic Programming are popular in the field of bioinformatics being extremely useful for some specific problems as DNA or amino acids sequences alignment, RNA structure prediction, protein structure research, and others (see [9], [10]).

In the case of sequence comparison analysis in Bioinformatics, a similarity measure on Σ together with a gap penalties function can be used to define the global similarity between two sequences in Σ^*. The computation is handled using the Needleman-Wunsch dynamic programming algorithm which is quite similar to the W-S-B algorithm for computation of distances. It is possible to define global similarity using a dynamic programming matrix.

To be more precise, let Σ be a non-empty set. Then a *free monoid* Σ^* on Σ is the monoid whose elements are all *finite* sequences of zero or more elements, from Σ with operation of *concatenation*. The set $\Sigma = \{A, B, C, ..., Z\}$ is called *alphabet*, and its elements $A, B, C, ..., Z$ are called **letters** of the alphabet, or *generators*. The elements $u \in \Sigma^*$ are called *words* or *strings*. The unique sequence of zero letters (the empty word) denoted by e is the *identity element* in Σ^*.

The *free semigroup* Σ^+ on Σ is defined as $\Sigma^+ := \Sigma^* \setminus \{e\}$.

Remark 4.1. Biological motivation

The macromolecules that contain the essential information of living cells can be represented as a family of words over a finite alphabet.

For example :

- **DNA molecules** are words in the free semigroup generated by the four letters **nucleotide alphabet** $\Sigma = \{A, C, T, G\}$.
- **RNA molecules** are words in the free semigroup generated by the four letters **nucleotide alphabet** $\Sigma = \{A, C, U, G\}$.
- **proteins molecules** can be regarded as words in the free semigroup whose generators are the 20 amino acids that compose the proteins in living cells **aminoacids alphabet** Σ_{AA}.

As an example, we mention here the insulin, whose intensive research, starting around 1950, has facilitated the development of the theory of molecular evolution of living organisms.

Insulin is present in almost all living organisms on the Earth, hence by comparing the insulin sequences found in different species and computing their similarity, one can get a very detailed insight into the evolution of life on Earth. Sequence comparison, similarity estimation, and so on, is one of the most fundamental research topics in bioinformatics (see [5]).

We define global similarity using a dynamic programming matrix.

Definition 4.2. ([10]) Let Σ be a set, $x, y \in \Sigma^*, s : \Sigma \times \Sigma \to \mathbb{R}$ and $g, h : \mathbb{N}^+ \to \mathbb{R}^+$. Let $m = |x|$ and $n = |y|$. The *Needleman-Wunsch* dynamic programming matrix, denoted NW(x, y, s, g, h), is an $(m+1) \times (n+1)$ matrix S with rows and columns indexed from 0 such that $S_{0,0} = 0, S_{i,0} = \max_{1 \le k \le i}\{S_{i-k,0} - h(k)\}, S_{0,j} = \max_{1 \le k \le j}\{S_{0,j-k} - g(k)\}$ and for all $i = 1, 2, ..., m$ and $j = 1, 2, ..., n$

$$S_{i,j} = \max\left\{S_{i-1,j-1} + s(x_i, y_j), \max_{1 \le k \le i}\{S_{i-k,j} - h(k)\}, \max_{1 \le k \le j}\{S_{i,j-k} - g(k)\}\right\}.$$

We define the *global similarity* between the sequences x and y (given s, g, and h), denoted $S(x, y)$, to be the value $S_{m,n}$.

Example 4.3. Let Σ be the alphabet, let $x = CTTCA$ and $y = CTACA$. For all $a, b \in \Sigma$, set $s(a, a) = 5, s(a, b) = -1$ if $a \neq b$ and let $g(k) = h(k) = 2k$. The matrix(or table) S used for computation of the Needleman-Wunsch distance is given in the Table 4.1.

The dynamic programming table used to compute the Needleman-Wunsch distance between the sequences CTTCA and CTACA. We can see that the global similarity between the sequences to be the value $S_{m,n} = S_{5,5} = 19$ and the path shown in bold in the Table 4.1 corresponds to the following alignment:

$$CTTCA$$

$$CTACA$$

Table 4.1: The Needleman-Wunsch dynamic programming table of example 4.3.

		0	1 C	2 T	3 T	4 C	5 A
0		0	-2	-4	-6	-8	-10
1	C	-2	**5**	3	1	-1	-3
2	T	-4	3	**10**	8	6	4
3	A	-6	1	8	**9**	7	11
4	C	-8	-1	6	7	**14**	12
5	A	-10	-3	4	5	12	**19**

Example 4.4. Let Σ be the alphabet, let $x = ATGGCGT$ and $y = ATGAGT$. For all $a, b \in \Sigma$, suppose $s(a,a) = 1, s(a,b) = 0$ if $a \neq b$ and let $g(k) = h(k) = 0$. The matrix(or table) S used for computation of the Needleman-Wunsch distance is given in the Table 4.2 observe that $S(x,y) = S_{6,7} = 5$.

Table 4.2: The Needleman-Wunsch dynamic programming table of example 4.4.

		0	1 A	2 T	3 G	4 G	5 C	6 G	7 T
0		0	0	0	0	0	0	0	0
1	A	0	1	1	1	1	1	1	1
2	T	0	1	2	2	2	2	2	2
3	G	0	1	2	3	3	3	3	3
4	A	0	1	2	3	3	3	3	3
5	G	0	1	2	3	4	4	4	4
6	T	0	1	2	3	4	4	4	**5**

The dynamic programming table used to compute the Needleman-Wunsch distance between the sequences ATGGCGT and ATGAGT. We can see that the global similarity between the sequences to be the value $S_{m,n} = S_{6,7} = 5$.

The global similarity $S(x,y)$ between sequences x and y defined above satisfies all conditions in the definition of similarity(see [10]).

Based on our Theorem 3.5 we can think as follows. Let us consider a finite set Π of biological sequences like for instance the insulin sequences in all species one can find in the NCBI data base. Clearly, this is a finite set of sequences, a finite set of data that can be considered as a compact set in Σ^*. By using dynamic programming, we can endow this compact set with a symmetric similarity function $S(x,y)$. Theorem 3.5 implies that there is always a Randers type metric whose associated distance function coincides with the weighted quasi distance function obtained from the similarity function.

We can conclude a new theory of research as follows:

Theorem 4.5. Let Π be a finite set of biological sequences and let $S(x, y)$ be the global similarity function given by the Needleman-Wunsch dynamic programming algorithm. If the associated distance ρ is an upper and lower curvature bounded, then there exists a metric of Randers type whose distance function coincides with the weighted quasi-metric induced by $S(x, y)$.

Chapter 5

Conclusions and Suggestions

5.1 Conclusions

In this research, we introduce definitions of the similarity function,metric space, quasi-metric space and weighted quasi-metric space. show that every weighted quasi-metric space can be identified with a subspace of a space of some canonical type, which is constructed from a metric space and present a very simple method to construct a weighted quasi-metric space, as the graph of function defined on metric space, and show that every weighted quasi-metric space arises in this way. We investigate a relation between weighted quasi-metric and similarity function, and we obtain a weighted quasi-metric space induced by a similarity function. On the other hand a similarity function can be induced by weighted quasi-metric space.

We study the geometrical properties of a topological space endowed with a similarity. We show that every space with a similarity function can be viewed as the graph of a 1-Lipschitz function over a suitable metric space and show that every similarity space can be constructed in this way. We discuss its relation with weighted quasi-metrics and Finsler metric of Randers type. Moreover, we present the relation of the mathematical concepts with computer science and bioinformatics. In conclusion, there is always a Randers type metric whose associated distance function coincides with the weighted quasi-metric induced by a similarity function.

5.2 Suggestions

Our research suggests that more study concerning weighted quasi-metric with similarity function and Finsler metric is necessary.

Another idea suggested by our research would be applied in Bioinformatics and computer science such as development of geometric aspects of the theory of weighted quasi-metric spaces and applications to similarity search in general and large protein datasets in particular.

References

[1] Bao, D., Chern, S.S. and Shen, Z., 2020. **An Introduction to Riemann-Finsler Geometry**. New York: Springer-Verlag.

[2] Chern, S.S. and Shen, Z., 2005. **Riemann-Finsler Geometry**. Singapore: World Scientific.

[3] Munkres, J.R., 2000. **Topology Second Edition**. Prentice Hall, Inc. Upper Saddle River, NJ07458, 2nd edition.

[4] Pevsner, J., 2017. **Bioinformatics And Functional Genomics**. Wiley India, 3rd edition.

[5] Sabau, S.V., 2015. **Introduction to metric and Length Spaces**. Tokai Univwesity.

[6] Sabau, S.V., Shibuya, K. and Shimada, H., 2014. **Metric structures associated to Finsler metric**. Publicationes Mathematica Debrecen, 84(1-2), 89-103.

[7] Shanker, G. and Rani, S., 2018. **Weighted quasi-metrics associated with Finsler metrics**. arXiv:1801.05636[math.DG].

[8] Shen, Z., 2001. **Lectures in Finsler Geometry**. World Scientific.

[9] Stojmirovic, A., 2004, **Quasi-metric space with measure**. Topology Proceedings, Vol.28, No.2, 655-671.

[10] Stojmirovic, A., 2008, **Quasi-metric, Similarities and Searches : aspects of geometry of protein datasets**. arXiv:0810.5407v1[cs.IR].

[11] Vitolo, P., 1995, **A representation theorem for quasi-metric Spaces**. Topology and its Apply, 65, 101-104.

[12] Vitolo, P., 1999, **The representation of weighted quasi-metric spaces**. Rendiconti dell'Istituto di Mathematica dell'Universita di Trieste, XXXI, 95-100.

Appendices

Current Applied Science and Technology Vol. 20 No. 2 (May-August 2020)

Finsler Metrics Induced by a Similarity Function

Nisachon Kumankat[1]*, Praiboon Pantaragphong[1] and Sorin V. Sabau[2]

[1] Department of Mathematics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
[2] School of Biological Science, Department of Biology, Tokai University, Sapporo, Japan

Received: 2 December 2019, Revised: 21 March 2020, Accepted: 2 April 2020

Abstract

In the present paper, the geometrical properties of a topological space endowed with a similarity was studied. Its relation with weighted quasi-metrics and Finsler metrics of Randers type was discussed. Finally, some applications to bioinformatics and computer science by relating similarities to dynamic programming algorithms are considered. In conclusion, the space containing the real-world data is non-symmetric and non-linear.

Keywords: Finsler metrics, similarity function, weighted quasi-metrics
DOI 10.14456/cast.2020.20

1. Introduction

Metric spaces, symmetric distances are used in different fields of pure mathematics like analysis, geometry and so on, as well as in applied mathematics, for instance in computer science, bioinformatics, data analysis, etc. These are the natural generalization of Euclidean and Riemannian spaces. On the other hand, non-symmetric distances, Minkowski norms, and Finsler metrics are also widely used in the analysis, differential geometry, data analysis, etc. We claim that symmetric distances, Euclidean and Riemannian metrics are just convenient approximations (usually obtained by averaging) of the real world. The real world, based on real data measurements is highly non-symmetric and non-linear. Of course, proving such a fact in its most generality is a very complex and difficult task, beyond the purpose of this paper.

However, we will argue that the similarity induced by the dynamic programming algorithm Needleman-Wunsch is actually equivalent in nature to non-symmetric distances (so-called quasi-distances) and Finsler metrics. Our main statement in this paper is that the following motions are equivalent in nature, i.e. symmetric similarity function, weighted quasi-distance and Finsler metrics of Randers type with reversible geodesics.

*Corresponding author: Tel.: +66 91 8517097
E-mail: Nisachon.Kumankat@gmail.com

The topic studied in the present paper is very important for analysis, geometry, computer science, data analysis, bioinformatics and so on because in some sense it shows that at least partially, the reality we are living in, is non-symmetric, non-linear, non-homogeneous and the study of data sets from the real world is actually equivalent to the use of weighted quasi-metrics on topological spaces, or of a Finsler metric of Randers type on smooth manifolds.

2. Similarities and Distances

We start by recalling the following definition [1, 2, 3].

Definition 2.1 Let X be a topological space. If $s : X \times X \to \mathbb{R}$ is a continuous mapping such that

i. $s(x,x) > 0$ for any $x \in X$,
ii. $s(x,x) \geq s(x,y)$ for any $x,y \in X$,
iii. if $s(x,y) = s(x,x)$ and $s(y,x) = s(y,y)$, then $x = y$ for any $x,y \in X$,
iv. $s(x,y) + s(y,z) \leq s(x,z) + s(y,y)$ for any $,y,z \in X$,

Then s is called a similarity function on X.

The relation with quasi-metrics is well-known [4, 5].

Definition 2.2 Let X be a non-empty set and d a real-valued function $d : X \times X \to [0,\infty)$ that satisfies:

i. $d(x,y) \geq 0$ and $d(x,y) = 0$ if and only if $x = y$ for any $x,y \in X$,
ii. $d(x,y) \leq d(x,z) + d(z,y)$ for any $x,y,z \in X$,
iii. if $d(x,y) = d(y,x) = 0$ then $x = y$ for any $x,y \in X$.

Then (X,d) is called a quasi-metric space.

Definition 2.3 A weighted quasi-metric space is a triple (X,d,w), where X is a non-emply set, $d : X \times X \to [0,\infty)$ and $w : X \to [0,\infty)$ that satisfies:

i. $d(x,y) \geq 0$ and $d(x,y) = 0$ if and only if $x = y$ for any $x,y \in X$,
ii. $d(x,y) \leq d(x,z) + d(z,y)$ for any $x,y,z \in X$,
iii. if $d(x,y) = d(y,x) = 0$ then $x = y$ for any $x,y \in X$,
iv. $d(x,y) + w(x) = d(y,x) + w(y)$ for any $x,y,z \in X$.

The function d is called a quasi-metric, and w is the weight function.

Proposition 2.4 If $s : X \times X \to \mathbb{R}$ be a similarity function on X, then $d : X \times X \to \mathbb{R}$ defined by

$$d(x,y) := s(y,y) - s(y,x), \qquad \text{for all } x,y \in X,$$

is a quasi-metric on X.

Proof Let $x,y,z \in X$. We verify the conditions in the definition of the quasi-metric.

i. Positiveness: $d(x,y) = s(y,y) - s(y,x)$,
Since $s(y,y) - s(y,x) \geq 0$, it is clear that $d(x,y) \geq 0$.
And $d(y,y) = s(y,y) - s(y,y) = 0$, then $x = y$.

ii. Triangle inequality:

$$\begin{aligned} d(x,y) &= s(y,y)-s(y,x) \\ &= s(y,y)-s(y,z)+s(y,z)-s(y,x) \\ &\leq s(y,y)-s(y,z)+s(z,z)-s(z,x) \\ &= d(z,y)+d(x,z) \\ &= d(x,z)+d(z,y). \end{aligned}$$

iii. Separation axiom:

$$d(x,y)=0 \ \wedge \ d(y,x)=0 \ \Rightarrow x=y$$
$$s(y,y)-s(y,x)=0 \ \wedge \ s(x,x)-s(x,y)=0 \ \Rightarrow x=y$$
$$s(y,y)=s(y,x) \ \wedge \ s(x,x)=s(x,y) \ \Rightarrow x=y.$$

Therefore, (X,d) is a quasi-metric on X.

Proposition 2.5 Let $s : X \times X \to \mathbb{R}$ is a similarity function on X. If s is a symmetric, i.e. $s(x,y)=s(y,x)$ for all $x,y \in X$, then (X,d) is a weighted quasi-metric space with weight function $w : X \to \mathbb{R}$, $w(x)=s(x,x)$.

Proof By proposition 2.4, we have $d(x,y)=s(y,y)-s(y,x)$.
Let $x,y,z \in X$. We verify the conditions in the definition of a weighted quasi-metric.

i. Positiveness: $d(x,y)=s(y,y)-s(y,x)$,
since $s(y,y)-s(y,x) \geq 0$, it is clear that $d(x,y) \geq 0$.
And $d(y,y)=s(y,y)-s(y,y)=0$, then $x=y$.

ii. Triangle inequality:

$$\begin{aligned} d(x,y) &= s(y,y)-s(y,x) \\ &= s(y,y)-s(y,z)+s(y,z)-s(y,x) \\ &\leq s(y,y)-s(y,z)+s(z,z)-s(z,x) \\ &= d(x,z)+d(z,y). \end{aligned}$$

iii. Separation axiom:

$$d(x,y)=0 \ \wedge \ d(y,x)=0 \ \Rightarrow x=y$$
$$s(y,y)-s(y,x)=0 \ \wedge \ s(x,x)-s(x,y)=0 \ \Rightarrow x=y$$
$$s(y,y)=s(y,x) \ \wedge \ s(x,x)=s(x,y) \ \Rightarrow x=y.$$

iv. Let $w : X \to \mathbb{R}$, $w(x)=s(x,x)$, we have

$$\begin{aligned} d(x,y)+w(x) &= s(y,y)-s(y,x)+s(x,x) \\ &= s(x,x)-s(x,y)+s(y,y) \\ &= d(y,x)+w(y). \end{aligned}$$

Therefore, (X,d) is a weighted quasi-metric space.

We will consider in the following only symmetric similarity function.

Remark 2.6 :

1) Observe that the quasi-distance d and the weight function w are determined only by the similarity function s.

2) Let we assume the quasi-distance d is actually a distance function, i.e. $d(x,y) = d(y,x)$ for all $x, y \in X$, it follows

$$s(y,y) - s(y,x) = s(x,x) - s(x,y),$$

and if we take into account that s is symmetric, then we obtain

$$s(x,x) = s(y,y), \qquad \text{for all } x, y \in X.$$

In other words, the quasi-distance induced by a similarity function is a distance if and only if the similarity is the same on the diagonal.

3) The symmetrized distance induced by a quasi-distance d is

$$\rho(x,y) := \frac{1}{2}[d(x,y) + d(y,x)] = \frac{1}{2}[s(x,x) + s(y,y)] - s(x,y).$$

Conversely, a weighted quasi-metric space induces a symmetric similarity function. Indeed, we have the following proposition hold

Proposition 2.7 Let (X, d) is a weighted quasi-metric space with the weight function $w : X \to \mathbb{R}$. Then the mapping $s : X \times X \to \mathbb{R}$,

$$s(x,y) := w(x) - d(y,x), \qquad \text{for all } x, y \in X,$$

is a symmetric similarity function on X.

Proof Let $x, y, z \in X$. We verify the conditions in definition 2.1, we have:

i. Since $w : X \to \mathbb{R}$ is weight function, then $w(x) > 0$ and $d(x,x) = 0$.
It is clear that $s(x,x) = w(x) - d(x,x) = w(x) > 0$.

ii. We will show that $s(x,x) - s(x,y) \geq 0$, we have

$$\begin{aligned} s(x,x) - s(x,y) &= w(x) - d(x,x) - w(x) + d(y,x) \\ &= -d(x,x) + d(y,x) \\ &= d(y,x). \end{aligned}$$

Since $d(y,x) \geq 0$, thus $s(x,x) - s(x,y) \geq 0$, i.e. $s(x,x) \geq s(x,y)$.

iii. Suppose that $s(x,y) = s(x,x)$ and $s(y,x) = s(y,y)$, we have

$$\begin{gathered} s(x,y) = s(x,x) \ \wedge \ s(y,x) = s(y,y) \ \Rightarrow \ x = y \\ w(x) - d(y,x) = w(x) - d(x,x) \ \wedge \ w(y) - d(x,y) = w(y) - d(y,y) \ \Rightarrow \ x = y \\ d(x,x) - d(y,x) = w(x) - w(x) \ \wedge \ d(y,y) - d(x,y) = w(y) - w(y) \ \Rightarrow \ x = y \\ d(x,x) - d(y,x) = 0 \ \wedge \ d(y,y) - d(x,y) = 0 \ \Rightarrow \ x = y \\ d(x,x) = d(y,x) \ \wedge \ d(y,y) = d(x,y) \ \Rightarrow \ x = y. \end{gathered}$$

iv. We will show that $s(x,y) + s(y,z) \leq s(x,z) + s(y,y)$, we have

$$\begin{aligned} s(x,y) + s(y,z) &= w(x) - d(y,x) + w(y) - d(z,y) \\ &= w(x) + w(y) - d(z,y) - d(y,x) \\ &\leq w(x) + w(y) - d(z,x) \\ &= w(x) - d(z,x) + w(y) - 0 \\ &= w(x) - d(z,x) + w(y) - d(y,y) \\ &= s(x,z) + s(y,y). \end{aligned}$$

Therefore, (X, s) is a symmetric similarity function on X.

Example 2.8 Let us consider the metric space (S, ρ) and the interval $I := (0, \infty)$. It is known that the product space $G := S \times I$ inherits a natural structure of generalized weighted quasi-metric structure (G, Q, W), where

$$Q : G \times G \to I, \quad Q(u,v) := \rho(x,y) + \eta - \xi,$$

$$W : G \to I, \qquad W(u) := 2\xi,$$

for any $u = (x, \xi),\ v = (y, \eta)$ on $G = S \times I$.

The similarity function $\varphi : G \times G \to \mathbb{R}$ induced by the weighted quasi-metric structure (G, Q, W) is given by

$$\varphi(u, v) := -\rho(x, y) + \xi + \eta, \qquad \text{for any } u = (x, \xi),\ v = (y, \eta) \in G.$$

Clearly this is a symmetric similarity function on G.

Indeed, let $u = (x, \xi),\ v = (y, \eta),\ l = (z, \zeta) \in G$, then

i. It is clear that

$$\varphi(u, u) = \varphi\big((x, \xi), (x, \xi)\big) = -\rho(x, x) + \xi + \xi = 2\xi > 0.$$

ii. We will show that $\varphi(u, u) - \varphi(u, v) \geq 0$, we have

$$\begin{aligned}
\varphi(u, u) - \varphi(u, v) &= -\rho(x, x) + \xi + \xi + \rho(x, y) - \xi - \eta \\
&= -\rho(x, x) + \rho(x, y) + \xi - \eta \\
&= \rho(x, y) + \xi - \eta \\
&= \rho(y, x) + \xi - \eta \\
&= Q\big((y, \eta), (x, \xi)\big) = Q(v, u) \quad \geq 0.
\end{aligned}$$

So, $\varphi(u, u) - \varphi(u, v) \geq 0$.

iii. Suppose that $\varphi(u, v) = \varphi(u, u)$ and $\varphi(v, u) = \varphi(v, v)$, that is

$$\begin{aligned}
\varphi\big((x, \xi), (y, \eta)\big) = \varphi\big((x, \xi), (x, \xi)\big) \ &\wedge\ \varphi\big((y, \eta), (x, \xi)\big) = \varphi\big((y, \eta), (y, \eta)\big) \\
-\rho(x, y) + \xi + \eta = -\rho(x, x) + \xi + \xi \ &\wedge\ -\rho(y, x) + \eta + \xi = -\rho(y, y) + \eta + \eta \\
\rho(x, x) - \rho(x, y) + \eta - \xi = 0 \ &\wedge\ \rho(y, y) - \rho(y, x) + \xi - \eta = 0,
\end{aligned}$$

by subtracting these two equalities we get, $\rho(x, x) - 2\xi = \rho(y, y) - 2\eta$, whence $x = y$, so that $\xi = \eta$. Hence $u = v$.

iv. We will show that $\varphi(u, v) + \varphi(v, l) \leq \varphi(u, l) + \varphi(v, v)$, we have

$$\begin{aligned}
\varphi(u, v) + \varphi(v, l) &= \varphi\big((x, \xi), (y, \eta)\big) + \varphi\big((y, \eta), (z, \zeta)\big) \\
&= -\rho(x, y) + \xi + \eta - \rho(y, z) + \eta + \zeta \\
&= -\rho(x, y) - \rho(y, z) + \xi + \zeta + \eta + \eta \\
&\leq -\rho(x, z) + \xi + \zeta + \eta + \eta \\
&= -\rho(x, z) + \xi + \zeta - \rho(y, y) + \eta + \eta \\
&= \varphi\big((x, \xi), (z, \zeta)\big) + \varphi\big((y, \eta), (y, \eta)\big) \\
&= \varphi(u, l) + \varphi(v, v).
\end{aligned}$$

Therefore, φ is a symmetric similarity function on G.

Example 2.9 Let (S, ρ) be a metric space and $f : S \to (0, \infty)$ a Lipschitz function with respect to ρ. Then it is known that the graph of f, i.e. $G_f = \{(x, f(x)) : x \in S\}$ has a weighted quasi-metric space structure (G_f, Q, W) given by

$$\begin{aligned}
&Q : G_f \times G_f \to (0, \infty), \quad Q(u, v) := \rho(x, y) + f(y) - f(x), \\
&W : G_f \to (0, \infty), \qquad W(u) := 2f(x),
\end{aligned}$$

for any $u = (x, f(x)),\ v = (y, f(y))$ on G_f.

It follows that the function $\varphi_f :\ G_f \times G_f \to \mathbb{R}$ given by

$$\varphi_f(u, v) := -\rho(x, y) + f(x) + f(y),$$

for any $u = (x, f(x)),\ v = (y, f(y)) \in G_f$, is a symmetric similarity function on G_f.

We can conclude that a metric space (X, ρ) with a Lipschitz function $f : X \to \mathbb{R}$ induces a similarity function on X.

The similarity space (G_f, φ_f) constructed here is called the bundle over the metric space (S, ρ).

3. Embeddings and relation to Finsler space

Let (X, σ) and (Y, τ) be two topoltgical spaces with similarities functions σ and τ, respectively. A continuous mapping $\varphi : X \to Y$ is called a similarity embedding if

$$\tau(\varphi(x), \varphi(y)) = \sigma(x, y),$$

for all $x, y \in X$.

Proposition 3.1 Let (X, q, w) and (Y, p, u) be two weighted quasi-metric spaces with the associated similarities σ and τ, respectively. The continuous function $\varphi : X \to Y$ is a similarity embedding if and only if it is an embedding of weighted quasi-metric spaces.

Proof We assume that $\varphi : (X, \sigma) \to (Y, \tau)$ is a similarity embedding, i.e.

$$\tau(\varphi(x), \varphi(y)) = \sigma(x, y), \qquad \forall x, y \in X.$$

The weighted quasi-metric (d, w) associated with the similarity function σ on X is given by

$$\begin{aligned} d(x, y) &= \sigma(y, y) - \sigma(y, x), && \forall x, y \in X, \\ w(x) &= \sigma(x, x), && \forall x \in X. \end{aligned}$$

The weighted quasi-metric $(\hat{d}, \widehat{w})$ associated with the similarity function τ on Y is given by

$$\begin{aligned} \hat{d}(x, y) &= \tau(y, y) - \tau(y, x), && \forall x, y \in Y, \\ \widehat{w}(x) &= \tau(x, x), && \forall x \in Y. \end{aligned}$$

We compute

$$\begin{aligned} \hat{d}(\varphi(x), \varphi(y)) &= \tau(\varphi(y), \varphi(y)) - \tau(\varphi(y), \varphi(x)) \\ &= \sigma(y, y) - \sigma(y, x) \\ &= d(x, y), \end{aligned}$$

for all $x, y \in X$. Likewise,

$$\widehat{w}(\varphi(x)) = \tau(\varphi(x), \varphi(x)) = \sigma(x, x) = w(x),$$

and hence it results that $\varphi : (X, d, w) \to (Y, \hat{d}, \widehat{w})$ is an embedding of weighted quasi-metrics spaces.

Conversely, we assume that $\varphi : (X, d, w) \to (Y, \hat{d}, \widehat{w})$ is an embedding of weighted quasi-metrics spaces, i.e.

$$\begin{aligned} \hat{d}(\varphi(x), \varphi(y)) &= d(x, y), \\ \widehat{w}(\varphi(x)) &= w(x), \end{aligned}$$

for all $x, y \in X$. Using now relations

$$\begin{aligned} \sigma(x, y) &= w(x) - d(y, x), && \forall x, y \in X, \\ \tau(x, y) &= \widehat{w}(x) - \hat{d}(y, x), && \forall x, y \in Y. \end{aligned}$$

We have

$$\begin{aligned} \tau(\varphi(x), \varphi(y)) &= \widehat{w}(\varphi(x)) - \hat{d}(\varphi(y), \varphi(x)) \\ &= w(x) - d(y, x) \\ &= \sigma(x, y), \end{aligned}$$

for all $x, y \in X$, therefore $\varphi : (X, \sigma) \to (Y, \tau)$ is a similarity embedding.

Theorem 3.2 Every space with a symmetry function (X, s) is embeddable in a bundle over a suitable metric space (S, ρ).

Proof The proof is quite straightforward by taking into account the construction in Example 2.9. This result can also be proved directly, using the fact that any weighted quasi-metric space is embeddable in a bundle over a suitable metric space [6, 7].

Theorem 3.3 1. Let (S, ρ) be a metric space and $f : S \to [0, \infty)$ a Lipschitz function on this metric space. Then the graph of f admits a similarity function $\varphi : G_f \times G_f \to \mathbb{R}$ that depends on ρ and f only.

2. Conversely, every similarity space (X, s) can be constructed in this way.

Proof (1) Statement 1 follows immediately from Example 2.9.

(2) Conversely, if we start with a similarity space (X, s), then we can consider :

(a) the associated weighted quasi-metric space (X, d, w), where d and w are given in Proposition 2.4 and 2.5.

(b) the symmetrized associated distance s has given in Remark 2.6.

By putting $f := \frac{1}{2} w$, using (X, s) and f, the construction from Statement 1 given that (G_f, φ) is a similarity space.

Moreover, (X, s) can be embedded in (G_f, φ) and the conclusion follows.

Lemma 3.4 Let M be a compact smooth manifold. If (M, ρ) is an upper and lower curvature bounded metric space, then there exists a Riemannian metric g on M whose distance function coincides with p.

Proof The proof is quite obvious. Every an upper and lower curvature bounded metric space (M, ρ) constructed on a n -dimensional compact smooth manifold M can be embedded isometrically in the Euclidean space $\mathbb{R}^{2n+1}$ with the canonical metric. On the other hand, the manifold M as a submanifold in $\mathbb{R}^{2n+1}$ inherits a canonical Riemannian metric [8, 9] from the embedding in the Euclidean space whose distance function obviously coincides with ρ.

Theorem 3.5 1. If $(M, F = \alpha + \beta)$ is a simply connected Randers defined by a Riemannian metric $\alpha = \sqrt{a_{ij}(x) y^i y^j}$ and a closed 1-form β, then M is endowed with a naturally induced similarity function.

2. Conversely, let s be a symmetric, similarity function defined on a compact differentiable manifold M whose associated distance p is upper and lower curvature bounded. If s is differentiable, then there exists a naturally constructed Randers metric on M that depends on s only.

Proof 1. It is clear since a Randers metric with β closed induces a weighted quasi metric on M.

2.Conversely, the similarity metric induces a weighted quasi metric (M, d, w). From Lemma 3.4 we can see that there exists a Riemannian metric on M whose distance function is exactly ρ, and since s was assumed smooth we can define $\beta := dw$, where w is the weight induced by s.

4. Relation with Bioinformatics and Computer Science

In order to assess the application of this theory, we start by recalling that Dynamic Programming is, at the same time, a mathematical optimization method as well as an algorithmic method in computer science. Dynamic Programming originates in the research of R. Bellman in the 1950s and it was applied eventually in many fields of science like engineering, economics and others. In the majority cases, this method works by simplifying a much more complicated problem by dividing it into much easier small problems using a recursive way. It is known that if a problem in computer science can be solved optimally by dividing it into smaller problems and then recursively determine the optimal solutions to these small problems, then the original problem has an optimal substructure. The algorithms involving Dynamic Programming are popular in the field of bioinformatics being extremely useful for some specific problems as DNA or amino acids sequences alignment, RNA structure prediction, protein structure research, and others [4, 5].

In the case of sequence comparison analysis in Bioinformatics, a similarity measure on Σ together with a gap penalties function can be used to define the global similarity between two sequences in Σ^*. The computation is handled using the Needleman-Wunsch dynamic programming algorithm which is quite similar to the W-S-B algorithm for computation of distances. It is possible to define global similarity using a dynamic programming matrix.

To be more precise, let Σ be a non-empty set. Then a free monoid Σ^*on Σ is the monoid whose elements are all finite sequences of zero or more elements, from Σ with the operation of concatenation. The set $\Sigma = \{A, B, C, \ldots, Z\}$ is called alphabet, and its elements $A, B, C, \ldots, Z$ are called **letters** of the alphabet, or generators. The elements $u \in \Sigma^*$ are called words or strings. The unique sequence of zero letters (the empty word) denoted by e is the identity element in Σ^*.

The free semigroup Σ^+on Σ is defined as $\Sigma^+ := \Sigma^* \setminus \{e\}$.

Remark 4.1 Biological motivation
The macromolecule that contains the essential information of living cells can be represented as a family of words over a finite alphabet. Consequently, DNA (or RNA) molecules can be seen as long words in the free semigroup with the generators $\Sigma = \{A, C, T, G\}$ (nucleotide alphabet). Proteins molecules can be regarded as words in the free semigroup whose generators are the 20 amino acids which compose the proteins in living cells (aminoacids alphabet) Σ_{AA}.

As an example, we mention here the insulin, whose intensive research, starting around 1950, has facilitated the development of the theory of molecular evolution of living organisms. Insulin is present in almost all living organisms on the Earth, hence by comparing the insulin sequences found in different species and computing their similarity, one can get a very detailed insight into the evolution of life on Earth. Sequence comparison, similarity estimation and so on, is one of the most fundamental research topics in bioinformatics [6].

We define global similarity using a dynamic programming matrix.

Definition 4.2 Let Σ be a set, $x, y \in \Sigma^*$, $s : \Sigma \times \Sigma \to \mathbb{R}$ and $g, h : \mathbb{N}^+ \to \mathbb{R}^+$. Let $x, y \in \Sigma^*$ and let $m = |x|$ and $n = |y|$. The Needleman-Wunsch dynamic programming matrix denoted $\mathrm{NW}(x, y, s, g, h)$, is an $(m+1) \times (n+1)$ matrix S with rows and columns indexed from 0 such that $S_{0,0} = 0$, $S_{i,0} = \max\limits_{1 \le k \le i}\{S_{i-k,0} - h(k)\}$, $S_{0,j} = \max\limits_{1 \le k \le j}\{S_{0,j-k} - g(k)\}$ and for all $i = 1,2, \ldots, m$ and $j = 1,2, \ldots, n$

$$S_{i,j} = \max\left\{S_{i-1,j-1} + s(x_i, y_j),\ \max_{1\le k\le i}\{S_{i-k,j} - h(k)\},\ \max_{1\le k\le j}\{S_{i,j-k} - g(k)\}\right\}.$$

We define the global similarity between the sequences x and y (given s, g and h), denoted $S(x, y)$, to be the value $S_{m,n}$.

The global similarity $S(x, y)$ between sequences x and y defined above satisfies all conditions in the definition of similarity [4].

Theorem 4.3 Let Π be a finite set of biological sequences and let $S(x, y)$ be the global similarity function given by the Needleman-Wunsch dynamic programming algorithm. If the associated distance p is upper and lower curvature bounded, then there exists a metric of Randers type whose distance function coincides with the weighted quasi-metric induced by $S(x, y)$.

Proof Based on our Theorem 3.5 we can explain as follows. Let us consider a finite set Π of biological sequences like, for instance, the insulin sequences in all species one can find in the NCBI database. Clearly, this is a finite set of sequences, a finite set of data that can be considered as a compact set in Σ^*. By using the dynamic programming, we can endow this compact set with a symmetric similarity function $S(x, y)$. Theorem 3.5 implies that there is always a Randers type metric whose associated distance function coincides with the weighted quasi distance function obtained from the similarity function.

5. Conclusions

In this paper, we introduce a definition of the similarity function, quasi-metric space and weighted quasi-metric space. We also study the geometrical properties of a topological space endowed with a similarity. The relation with embeddings and bundle and Finsler metrics of Randers type has been explained. Moreover, we present the relation of the mathematical concepts with computer science and bioinformatics. In conclusion, there is always a Randers type metric whose associated distance function coincides with the weighted quasi-metric induced by a similarity function.

References

[1] Sabau, S.V., Shibuya, K. and Shimada, H., 2014. Metric structures associated to Finsler metrics. Publications Mathematicae Debrecen, 84(1-2), 89-103.

[2] Shanker, G. and Rani, S., 2018. Weighted quasi-metrics associated with Finsler metrics. [online] Available at : https://www.researchgate.net/publication/322567912_Weighted_quasi-metrics_associated_with_Finsler_metrics

[3] Vitolo, P., 1999. The representation of weighted quasi-metric spaces. Rendiconti dell'Istituto di Matematica dell'Università di Trieste, 31, 95-100.

[4] Stojmirovic, A., 2008. Quasi-metrics, Similarities and Searches: Aspects of Geometry of Protein Datasets. Ph.D. Victoria University of Wellington.

[5] Stojmirovic, A., 2004. Quasi-metric space with measure. Topology Proceedings, 28 (2), 655-671.

[6] Pevsner, J., 2017. Bioinformatics and Functional Genomics. 3rd ed. New Delhi: Wiley India.

[7] Vitolo, P., 1995. A representation theorem for quasi-metric spaces. Topology and Its Applications, 65, 101-104.

[8] Bao, D., Chern, S.S. and Shen, Z., 2000. An Introduction to Riemann-Finsler Geometry. New York: Springer-Verlag.
[9] Chern, S.S. and Shen, Z., 2005. Riemann-Finsler Geometry. 6th ed. Singapore: World Scientific.

Author Biography

Name	Miss Nisachon Kumankat
Date of Birth	4 June 1995
Address	122 m.7 Maetha, Lamphun, 51170
Education	(2017) Bachelor of Science in Applied Mathematics GPA 3.73 King Mongkut's Institute of Technology Ladkrabang
Scholarship	Scholarships for Graduate Student, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang.

Academic Publications

Kumankat, N., Pantaragphong, P. and Sabau, S.V., 2020. Finsler Metrics Induced by a Similarity Function. Current Applied Science and Technology, Volume 2, No.2, May-August 2020, page 334-343.