Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 942687, 7 pages
http://dx.doi.org/10.1155/2013/942687

Research Article

A Class of New Metrics for n-Dimensional Unit Hypercube

Guoxiang Lu

School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430077, China

Correspondence should be addressed to Guoxiang Lu; lgxmath@znufe.edu.cn

Received 30 May 2013; Revised 19 August 2013; Accepted 22 August 2013

Academic Editor: Mohamad Alwash

We study the metric on the n-dimensional unit hypercube. We introduce a class of new metrics for the space, which is information theoretically motivated and has close relation to Jensen-Shannon divergence. These metrics are obtained by discussing a function $FD_\alpha(P,Q)$ with the parameter α. We come to the conclusion that the sufficient and necessary condition of the function being a metric is $0<\alpha\leq 1/2$. Finally, by computing basic examples of codons, we show some numerical comparison of the new metrics to the former metric.

1. Introduction

As early as 1971, Zadeh introduced a geometric interpretation of fuzzy sets by stating that they can be represented as points in unit hypercube [1]. Many years later, his idea was taken up by Kosko, who built a promising fuzzy-theoretical framework and geometry thereon [2, 3]. This geometry of fuzzy sets was used in [4] to develop the fuzzy polynucleotide space. He demonstrated a polynucleotide molecule as a point in an n-dimensional unit hypercube. This approach enabled us to make quantitative studies such as the measurement of distances, similarities, and dissimilarities between polynucleotide sequences. The n-dimensional unit hypercube enriched by a metric d is named fuzzy polynucleotide space (I^n, d) with $I=[0,1]\subset\mathbb{R}$ which is a metric space. Torres and Nieto [5, 6] considered the frequencies of the nucleotides at the three base sites of a codon in the coding sequence as fuzzy sets to give an example on I^{12}. Later, Dress, Lokot, and Pustyl'nikov have pointed out that the metric is under the L_1-norm and showed the metric properties [7].

Because the fuzzy sets which come from the polynucleotide molecules reflect the information of those sequences, we may introduce the related concept in information theory to measure the differences between polynucleotide sequences. In information theory, the relative entropy (also called Kullback-Leibler divergence) is the most common measure to show two probability distributions. But it is not a metric for it does not satisfy symmetric and triangle inequality [8]. In the past time, many pieces of research [9–15] were made to improve the relative entropy. From those references, Jensen-Shannon divergence as an improvement of relative entropy received much attention. In this paper, a class of new metrics inspired by the Jensen-Shannon divergence are introduced in the n-dimensional unit hypercube. These metrics with information-theoretical property of logarithm can replace the former metric d in the fuzzy polynucleotide space.

2. Preliminaries

Let $X=\{x_1,x_2,\ldots,x_n\}$ be a fixed set; a fuzzy set in X is defined by

$$A=\{(x,\mu_A(x))\mid x\in X\}, \tag{1}$$

where

$$\mu_A : X\longrightarrow I=[0,1], \qquad x\longmapsto \mu_A(x). \tag{2}$$

The number $\mu_A(x)$ denotes the membership degree of the element x in the fuzzy set A. We can also use the unit hypercube $I^n=[0,1]^n$ to describe all the fuzzy sets in X, because a fuzzy set A determines a point $P=(\mu_A(x_1),\mu_A(x_2),\ldots,\mu_A(x_n))$. Reciprocally, any point $P=(a_1,a_2,\ldots,a_n)\in I^n$ generates a fuzzy set A defined by $\mu_A(x_i)=a_i$, $i=1,2,\ldots,n$. A polynucleotide is representable as such an ordered fuzzy set in [4–6]. Given two fuzzy sets

$P = (p_1, p_2, \ldots, p_n)$, $Q = (q_1, q_2, \ldots, q_n) \in I^n$, the metric d is defined by

$$d(P, Q) = \frac{\sum_{i=1}^{n} |p_i - q_i|}{\sum_{i=1}^{n} \max\{p_i, q_i\}}. \tag{3}$$

With the metric d defined, the fuzzy polynucleotide space is constructed.

Let Y be a discrete random variable with alphabet $\mathscr{Y} = \{y_1, y_2, \ldots, y_n\}$. P, Q are two probability distributions of Y. Then, the relative entropy between P and Q is defined as

$$D(P \parallel Q) = \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i}. \tag{4}$$

Here, ln denotes the natural logarithm for convenience. Furthermore, the Jensen-Shannon divergence is defined by

$$\operatorname{JSD}(P, Q) = \frac{1}{2}\left[D(P \parallel R) + D(Q \parallel R)\right], \tag{5}$$

where $R = (1/2)(P + Q)$.

Jensen-Shannon divergence is obviously nonnegative, symmetric and vanishes for $P = Q$, but it does not fulfill the triangle inequality. And a point in the n-dimensional I^n is not a probability distribution. In view of the foregoing, the concept of Jensen-Shannon divergence should be generalized. If P, Q are two points in I^n, this function $FD_\alpha(P, Q)$ is studied:

$$\begin{aligned} FD_\alpha(P, Q) &= (2\operatorname{JSD}(P, Q))^\alpha \\ &= \left[D(P \parallel R) + D(Q \parallel R)\right]^\alpha \\ &= \left(\sum_{i=1}^{n}\left(p_i \ln \frac{2p_i}{p_i + q_i} + q_i \ln \frac{2q_i}{p_i + q_i}\right)\right)^\alpha, \end{aligned} \tag{6}$$

where $\alpha \in \mathbb{R}$. In the following sections, we discuss the function to all $\alpha \in \mathbb{R}$ and obtain the class of new metrics.

3. Auxiliary Result Associated with $FD_\alpha(P,Q)$

Definition 1. Let the function $J(p, q) : [0, +\infty) \times [0, +\infty) \to \mathbb{R}$ be defined by

$$J(p, q) = p \ln \frac{2p}{p + q} + q \ln \frac{2q}{p + q}. \tag{7}$$

In the above definition, we use convention based on continuity that $0 \ln(0/x) = 0$, $x \in [0, +\infty)$.

To all $\alpha \in \mathbb{R}$, we wonder whether the function $(J(p, q))^\alpha$ can be a metric on the space I.

Lemma 2. *$J(p, q) \geq 0$, with equality only for $p = q$.*

Proof. From (7) we can get

$$\begin{aligned} &J(p, q) \\ &\quad = (p + q)\left(\frac{p}{p + q} \ln \frac{p/(p + q)}{1/2} + \frac{q}{p + q} \ln \frac{q/(p + q)}{1/2}\right). \end{aligned} \tag{8}$$

The formula above expresses that $J(p, q)$ is the relative entropy of the probability distributions $(p/(p + q), q/(p + q))$ and $(1/2, 1/2)$. With the nonnegativity of the relative entropy [8] the lemma holds. □

Lemma 3. *If the function $f : [0, +\infty) \to \mathbb{R}$ is defined by $f(x) = J(a, x) - a(\ln(2a/(a + x)))^2$ with $a > 0$, then f is convex function.*

Proof. Straightforward derivative shows

$$\begin{aligned} f'(x) &= \ln \frac{2x}{a + x} + \frac{2a}{a + x} \ln \frac{2a}{a + x}, \\ f''(x) &= \frac{a(a - x - 2x \ln(2a/(a + x)))}{x(a + x)^2}. \end{aligned} \tag{9}$$

Using the standard inequality

$$\ln a \geq 1 - \frac{1}{a}, \qquad \ln a \leq a - 1, \tag{10}$$

we find

$$\begin{aligned} a - x - 2x \ln \frac{2a}{a + x} &\geq a - x - 2x\left(\frac{2a}{a + x} - 1\right) \\ &= \frac{(a - x)^2}{a + x} \geq 0. \end{aligned} \tag{11}$$

The equality holds if and only if $x = a$. So $f''(x) \geq 0$, f is convex function, and the function $f(x)$ gets the minimum 0 when $x = a$ for $f'(a) = 0$. □

As a consequence of Lemma 3, when $x \neq a$,

$$\frac{1}{a} > \frac{(\ln(2a/(a + x)))^2}{J(a, x)}. \tag{12}$$

Lemma 4. *If the function $g : [0, +\infty) \to \mathbb{R}$ is defined by $g(x) = \ln(2x/(a + x))/\sqrt{J(a, x)}$ with $a > 0$, then*

$$\lim_{x \to a^+} g(x) = \sqrt{\frac{1}{a}}, \qquad \lim_{x \to a^-} g(x) = -\sqrt{\frac{1}{a}}. \tag{13}$$

Proof. As $g^2(x) = (\ln(2x/(a + x)))^2/J(a, x)$, using l'Hôspital's rule we can obtain

$$\lim_{x \to a} g^2(x) = \lim_{x \to a} \frac{2a \ln(2x/(a + x))/x(a + x)}{\ln(2x/(a + x))} = \frac{1}{a}. \tag{14}$$

And $g(x) < 0$ in the case $x < a$, $g(x) > 0$ in the case $x > a$. Thus, the lemma holds. □

Assuming $0 < p < q$, we introduce the function $h : [0, +\infty) \to [0, +\infty)$ defined by

$$h(r) = \sqrt{J(p, r)} + \sqrt{J(q, r)}. \tag{15}$$

Lemma 5. *The function $h(r)$ has two minima, one at $r = p$ and the other at $r = q$.*

Proof. The derivative of the function h is

$$h'(r) = \frac{1}{2}\left(\frac{\ln(2r/(p+r))}{\sqrt{J(p,r)}} + \frac{\ln(2r/(p+r))}{\sqrt{J(q,r)}}\right). \quad (16)$$

So $h'(r) < 0$ for $r \in [0,p)$ and $h'(r) > 0$ for $r \in (q,+\infty)$. It shows h is monotonic decreasing in $[0,p)$ and monotonic increasing in $[q,+\infty)$.

Next, consider the monotonicity of h in the open interval (p,q).

From (12), we have

$$\sqrt{\frac{1}{p}} > \frac{|\ln(2p/(p+q))|}{\sqrt{J(p,q)}},$$
$$\sqrt{\frac{1}{q}} > \frac{|\ln(2q/(p+q))|}{\sqrt{J(p,q)}}. \quad (17)$$

From Lemma 4, we have

$$\lim_{r\to p^+}\frac{|\ln(2r/(p+r))|}{\sqrt{J(p,r)}} = \sqrt{\frac{1}{p}},$$
$$\lim_{r\to q^-}\frac{|\ln(2r/(q+r))|}{\sqrt{J(q,r)}} = -\sqrt{\frac{1}{q}}. \quad (18)$$

Using (17) and (18), we obtain

$$\lim_{r\to p^+} h'(r) = \frac{1}{2}\left(\lim_{r\to p^+}\frac{\ln(2r/(p+r))}{\sqrt{J(p,r)}} + \frac{\ln(2p/(p+q))}{\sqrt{J(p,q)}}\right)$$
$$= \frac{1}{2}\left(\sqrt{\frac{1}{p}} + \frac{\ln(2p/(p+q))}{\sqrt{J(p,q)}}\right) > 0,$$
$$\lim_{r\to q^-} h'(r) = \frac{1}{2}\left(\frac{\ln(2q/(p+q))}{\sqrt{J(p,q)}} + \lim_{r\to q^-}\frac{\ln(2r/(q+r))}{\sqrt{J(q,r)}}\right)$$
$$= \frac{1}{2}\left(\frac{\ln(2q/(p+q))}{\sqrt{J(p,q)}} - \sqrt{\frac{1}{q}}\right) < 0. \quad (19)$$

Let

$$A(y,r) = \frac{\ln(2r/(y+r))}{\sqrt{J(y,r)}} = \frac{\ln(2r/(y+r))}{\sqrt{r}\sqrt{J(y/r,1)}} = \frac{1}{\sqrt{r}}B(y,r); \quad (20)$$

then

$$\frac{\partial B(y,r)}{\partial r} = \left(y\left(2r\ln\frac{2r}{y+r} + \ln\frac{2y}{y+r}\left(2y + (y+r)\ln\frac{2r}{y+r}\right)\right)\right) \times \left(2\sqrt{r}(y+r)(J(y,r))^{3/2}\right)^{-1} \triangleq \frac{yC(r)}{2\sqrt{r}(y+r)(J(y,r))^{3/2}}. \quad (21)$$

We have $C(y) = 0$, $C'(r) = \ln(2r/(y+r)) + (1/r)\ln(2y/(y+r))(y + r\ln(2r/(y+r)))$, $C'(y) = 0$, and $C''(r) = -(1/(r^2(y+r)))(y^2\ln(2y/(y+r)) + r^2\ln(2r/(y+r)))$. From (10),

$$y^2\ln\frac{2y}{y+r} + r^2\ln\frac{2r}{y+r} \geq y^2\left(1 - \frac{y+r}{2y}\right) + r^2\left(1 - \frac{y+r}{2r}\right) = \frac{(y-r)^2}{2} \geq 0. \quad (22)$$

The equality holds if and only if $r = y$. This means $C''(r) \leq 0$, and the equality holds if and only if $r = y$. From the above, $C(r) \leq 0$ is easily obtained, and $C(r) = 0$ if and only if $r = y$. So with respect to variable r in the open interval (p,q), $B(p,r)$ and $B(q,r)$ are both monotonic decreasing, and $B(p,r) + B(q,r)$ is also monotonic decreasing. As

$$h'(r) = \frac{1}{2}(A(p,r) + A(q,r)) = \frac{1}{2\sqrt{r}}(B(p,r) + B(q,r)), \quad (23)$$

$\lim_{r\to p^+} B(p,r)+B(q,r) > 0$, and $\lim_{r\to q^-} B(p,r)+B(q,r) < 0$, we can see that $B(p,r) + B(q,r)$ has only one zero point in the open interval (p,q) with respect to variable r. As a consequence, $h'(r)$ has only one zero point x_0 in the open interval (p,q) with respect to variable r. This means $h'(r) > 0$ in the interval (p,x_0), $h'(r) < 0$ in the interval (x_0,q). From this, we know h has only one maximum and no minimum in the open interval (p,q).

As a result, the conclusion in the lemma is obtained. □

Theorem 6. *The function $(J(p,q))^{1/2}$ is a metric on the space I.*

Proof. From Lemma 2, the function $J(p,q) \geq 0$ with equality only for $p = q$ is proved. Hence, $(J(p,q))^{1/2} \geq 0$ with equality only for $p = q$. It is easy to see that $(J(p,q))^{1/2} = (J(q,p))^{1/2}$. Because the formula holds,

$$(J(p,q))^{1/2} = \lim_{r\to p}(J(p,r))^{1/2} + (J(q,r))^{1/2}$$
$$= \lim_{r\to q}(J(p,r))^{1/2} + (J(q,r))^{1/2}, \quad (24)$$

and from Lemma 5, the triangle inequality

$$(J(p,q))^{1/2} \le (J(p,r))^{1/2} + (J(q,r))^{1/2} \tag{25}$$

can be easily proved for any number $r \in I$. □

Corollary 7. *If $0 < \alpha < 1/2$, then the function $(J(p,q))^{\alpha}$ is a metric on the space I.*

Proof. The properties of nonnegativity and symmetry can be proved using the same method in Theorem 6.

Let $a, b > 0$ and $0 < \gamma < 1$, then $a^{\gamma} + b^{\gamma} > (a+b)^{\gamma}$ which follows from the concavity of x^{γ}. Now a γ which satisfies $\alpha = (1/2)\gamma$ can be found. Thus,

$$\begin{aligned}(J(p,r))^{\alpha} + (J(q,r))^{\alpha} &= (J(p,r))^{(1/2)\gamma} + (J(q,r))^{(1/2)\gamma} \\ &\ge \left((J(p,r))^{1/2} + (J(q,r))^{1/2}\right)^{\gamma} \\ &\ge (J(p,q))^{(1/2)\gamma} = (J(p,q))^{\alpha}.\end{aligned} \tag{26}$$

This is the triangle inequality for the function $(J(p,q))^{\alpha}$. □

Theorem 8. *If $1/2 < \alpha < 1$, then function $(J(p,q))^{\alpha}$ is not a metric on the space I.*

Proof. Assuming $0 < p < q$, let $l(r) = (J(p,r))^{\alpha} + (J(q,r))^{\alpha}$. Firstly, the formula holds:

$$\begin{aligned}(J(p,q))^{\alpha} &= \lim_{r\to p}(J(p,r))^{\alpha} + (J(q,r))^{\alpha} \\ &= \lim_{r\to q}(J(p,r))^{\alpha} + (J(q,r))^{\alpha}.\end{aligned} \tag{27}$$

The derivative of the function l is

$$l'(r) = \alpha\left(\ln\frac{2r}{p+r}(J(p,r))^{\alpha-1} + \ln\frac{2r}{q+r}(J(q,r))^{\alpha-1}\right). \tag{28}$$

Let

$$\begin{aligned}m(r) &= \left(\ln\frac{2r}{p+r}(J(p,r))^{\alpha-1}\right)^{1/(1-\alpha)} \\ &= \frac{(\ln(2r/(p+r)))^{1/(1-\alpha)}}{J(p,r)}.\end{aligned} \tag{29}$$

Using l'Hôspital's rule,

$$\lim_{r\to p} m(r) = \frac{p}{r(p+r)(1-\alpha)}\left(\ln\frac{2r}{p+r}\right)^{(2\alpha-1)/(1-\alpha)} = 0. \tag{30}$$

So

$$\lim_{r\to p} l'(r) = \alpha\ln\frac{2p}{p+q}(J(p,q))^{\alpha-1} < 0. \tag{31}$$

According to the definition of derivative, there exists a $\delta > 0$ such that, for any $s \in (p, p+\delta)$,

$$\begin{aligned}(J(p,q))^{\alpha} &= \lim_{r\to p^{+}}(J(p,r))^{\alpha} + (J(q,r))^{\alpha} \\ &> (J(p,s))^{\alpha} + (J(q,s))^{\alpha}.\end{aligned} \tag{32}$$

This shows that the triangle inequality does not hold. □

Theorem 9. *If $\alpha \ge 1$, then function $(J(p,q))^{\alpha}$ is not a metric on the space I.*

Proof. Let $p, q \in I$ and $p \ne q$; $r = (p+q)/2$. Consider the following:

$$\begin{aligned}J(p,q) - J(p,r) - J(q,r) &= p\ln\frac{2p}{p+q} + q\ln\frac{2q}{p+q} \\ &\quad - \left(p\ln\frac{2p}{p+r} + r\ln\frac{2r}{p+r}\right) \\ &\quad - \left(r\ln\frac{2r}{q+r} + q\ln\frac{2q}{q+r}\right).\end{aligned} \tag{33}$$

By substituting $r = (p+q)/2$,

$$\begin{aligned}&J(p,q) - J(p,r) - J(q,r) \\ &\quad = \frac{3p+q}{2}\ln\frac{3p+q}{2p+2q} + \frac{p+3q}{2}\ln\frac{p+3q}{2p+2q} \\ &\quad = J\left(\frac{3p+q}{2}, \frac{p+3q}{2}\right) > 0.\end{aligned} \tag{34}$$

For $(3p+q)/2, (p+3q)/2 \in I$ and $(3p+q)/2 \ne (p+3q)/2$.

Because if $0 \le x \le 1$, then $x^{\alpha} \le x$ and $(1-x)^{\alpha} \le 1-x$, we have $x^{\alpha} + (1-x)^{\alpha} \le 1$. Thus,

$$\begin{aligned}&\left(\frac{J(p,r)}{J(p,q)}\right)^{\alpha} + \left(\frac{J(q,r)}{J(p,q)}\right)^{\alpha} \\ &\quad < \left(\frac{J(p,r)}{J(p,r)+J(q,r)}\right)^{\alpha} + \left(\frac{J(q,r)}{J(p,r)+J(q,r)}\right)^{\alpha} \\ &\quad \le \frac{J(p,r)}{J(p,r)+J(q,r)} + \frac{J(q,r)}{J(p,r)+J(q,r)} \\ &\quad = 1.\end{aligned} \tag{35}$$

As a consequence, $(J(p,q))^{\alpha} > (J(p,r))^{\alpha} + (J(q,r))^{\alpha}$. This shows that the triangle inequality does not hold. □

To sum up the theorems and corollary above, we can obtain the main theorem.

Theorem 10. *The function $(J(p,q))^{\alpha}$ is a metric on the space I if and only if $0 < \alpha \le 1/2$.*

4. Metric Property of $FD_\alpha(P,Q)$

In this section, we mainly prove the following theorem.

Theorem 11. *The function $FD_\alpha(P,Q)$ is a metric on the space I^n if and only if $0 < \alpha \le 1/2$.*

Proof. When $\alpha \le 0$, $D(P \parallel R) + D(Q \parallel R) = 0$, where $P = Q$. So the function $FD_\alpha(P,Q) = [D(P \parallel R) + D(Q \parallel R)]^\alpha$ is not a metric. When $\alpha > 0$, From (6), we can get $FD_\alpha(P,Q) = (\sum_{i=1}^n J(p_i,q_i))^\alpha$. It is easy to see that $FD_\alpha(P,Q) \ge 0$ with equality only for $P = Q$, and $FD_\alpha(P,Q) = FD_\alpha(Q,P)$. So what we are concerned with is whether the triangle inequality

$$FD_\alpha(P,Q) \le FD_\alpha(P,H) + FD_\alpha(Q,H) \tag{36}$$

holds for any $P, Q, H \in I^n$.

When $P = Q$, $FD_\alpha(P,Q) = 0$, the triangle inequality (36) holds apparently. So we assume $P \ne Q$ in the following.

Next consider the value of α in three cases, respectively,

(i) $0 < \alpha \le 1/2$.

From Theorem 10, the inequality $(J(p_i,q_i))^\alpha \le (J(p_i,h_i))^\alpha + (J(q_i,h_i))^\alpha$ holds. Applying Minkowski's inequality, we have

$$\begin{aligned}
\left(\sum_{i=1}^n J(p_i,q_i)\right)^\alpha &= \left\{\sum_{i=1}^n \left((J(p_i,q_i))^\alpha\right)^{1/\alpha}\right\}^\alpha \\
&\le \left\{\sum_{i=1}^n \left((J(p_i,h_i))^\alpha + (J(q_i,h_i))^\alpha\right)^{1/\alpha}\right\}^\alpha \\
&\le \left\{\sum_{i=1}^n \left((J(p_i,h_i))^\alpha\right)^{1/\alpha}\right\}^\alpha + \left\{\sum_{i=1}^n \left((J(q_i,h_i))^\alpha\right)^{1/\alpha}\right\}^\alpha \\
&= \left(\sum_{i=1}^n J(p_i,h_i)\right)^\alpha + \left(\sum_{i=1}^n J(q_i,h_i)\right)^\alpha.
\end{aligned} \tag{37}$$

So the triangle inequality (36) holds:

(ii) $\alpha \ge 1$.

Let $R = (1/2)(P + Q)$; then using Theorem 9 and noticing $P \ne Q$, we find

$$\begin{aligned}
&\frac{FD_\alpha(P,R) + FD_\alpha(Q,R)}{FD_\alpha(P,Q)} \\
&\quad = \frac{(\sum_{i=1}^n J(p_i,r_i))^\alpha + (\sum_{i=1}^n J(q_i,r_i))^\alpha}{(\sum_{i=1}^n J(p_i,q_i))^\alpha} \\
&\quad < \frac{(\sum_{i=1}^n J(p_i,r_i))^\alpha + (\sum_{i=1}^n J(q_i,r_i))^\alpha}{(\sum_{i=1}^n (J(p_i,r_i) + J(q_i,r_i)))^\alpha} \\
&\quad \le 1.
\end{aligned} \tag{38}$$

This means $FD_\alpha(P,Q) > FD_\alpha(P,R) + FD_\alpha(Q,R)$. The inequality is not consistent with triangle inequality (36):

(iii) $1/2 < \alpha < 1$.

Let

$$F(x_1,\ldots,x_n) = F_1(x_1,\ldots,x_n) + F_2(x_1,\ldots,x_n), \tag{39}$$

where

$$\begin{aligned}
F_1(x_1,\ldots,x_n) &= \left(\sum_{i=1}^n \left(p_i \ln \frac{2p_i}{p_i + x_i} + x_i \ln \frac{2x_i}{p_i + x_i}\right)\right)^\alpha, \\
F_2(x_1,\ldots,x_n) &= \left(\sum_{i=1}^n \left(x_i \ln \frac{2x_i}{q_i + x_i} + q_i \ln \frac{2q_i}{q_i + x_i}\right)\right)^\alpha.
\end{aligned} \tag{40}$$

Then, $F(p_1,\ldots,p_n) = F(q_1,\ldots,q_n) = FD_\alpha(P,Q)$.

Next, we prove $(p_1,\ldots,p_n)$ and $(q_1,\ldots,q_n)$ are not the extreme points of the function $F(x_1,\ldots,x_n)$. From the symmetry, we only need to prove $(p_1,\ldots,p_n)$ is not the extreme point.

By partial derivative,

$$\begin{aligned}
&\left.\frac{\partial F}{\partial x_i}\right|_{(x_1,\ldots,x_n)=(p_1,\ldots,p_n)} \\
&\quad = \left.\frac{\partial F_1}{\partial x_i}\right|_{(x_1,\ldots,x_n)=(p_1,\ldots,p_n)} + \left.\frac{\partial F_2}{\partial x_i}\right|_{(x_1,\ldots,x_n)=(p_1,\ldots,p_n)}.
\end{aligned} \tag{41}$$

Since $P \ne Q$, we might as well assume $p_1 \ne q_1$ and $p_1 > 0$. Consider the following

$$\begin{aligned}
&\left.\frac{\partial F_2}{\partial x_1}\right|_{(x_1,\ldots,x_n)=(p_1,\ldots,p_n)} \\
&\quad = \alpha \ln \frac{2p_1}{p_1 + q_1} \cdot \left(\sum_{i=1}^n \left(p_i \ln \frac{2p_i}{p_i + q_i} + q_i \ln \frac{2q_i}{p_i + q_i}\right)\right)^{\alpha-1} \\
&\quad \ne 0,
\end{aligned} \tag{42}$$

$$\begin{aligned}
&\left.\frac{\partial F_1}{\partial x_1}\right|_{(x_1,\ldots,x_n)=(p_1,\ldots,p_n)} \\
&\quad = \lim_{\Delta x_1 \to 0} \frac{1}{\Delta x_1} \left(F_1(p_1 + \Delta x_1,\ldots,p_n) - F_1(p_1,\ldots,p_n)\right) \\
&\quad = \lim_{\Delta x_1 \to 0} \frac{1}{\Delta x_1} \\
&\qquad \cdot \left(p_1 \ln \frac{2p_1}{2p_1 + \Delta x_1} + (p_1 + \Delta x_1) \ln \frac{2(p_1 + \Delta x_1)}{2p_1 + \Delta x_1}\right)^\alpha.
\end{aligned} \tag{43}$$

Using (10), we have

$$\ln \frac{2p_1}{2p_1+\Delta x_1} \le \frac{2p_1}{2p_1+\Delta x_1} - 1 = -\frac{\Delta x_1}{2p_1+\Delta x_1},$$
$$\ln \frac{2(p_1+\Delta x_1)}{2p_1+\Delta x_1} \le \frac{2(p_1+\Delta x_1)}{2p_1+\Delta x_1} - 1 = \frac{\Delta x_1}{2p_1+\Delta x_1}. \quad (44)$$

If Δx_1 is small enough, using (44), we find the inequality

$$\begin{aligned}&\left(p_1 \ln \frac{2p_1}{2p_1+\Delta x_1} + (p_1+\Delta x_1)\ln\frac{2(p_1+\Delta x_1)}{2p_1+\Delta x_1}\right)^{\alpha}\\ &\le \left(p_1\left(-\frac{\Delta x_1}{2p_1+\Delta x_1}\right) + (p_1+\Delta x_1)\left(\frac{\Delta x_1}{2p_1+\Delta x_1}\right)\right)^{\alpha}\\ &= \frac{(\Delta x_1)^{2\alpha}}{(2p_1+\Delta x_1)^{\alpha}}.\end{aligned} \quad (45)$$

A straight result of (45) is

$$\begin{aligned}0 &\le \left|\frac{1}{\Delta x_1}\right| \\ &\quad\cdot \left|\left(p_1 \ln \frac{2p_1}{2p_1+\Delta x_1} + (p_1+\Delta x_1)\ln\frac{2(p_1+\Delta x_1)}{2p_1+\Delta x_1}\right)^{\alpha}\right| \\ &\le \frac{|\Delta x_1|^{2\alpha-1}}{(2p_1+\Delta x_1)^{\alpha}} \longrightarrow 0, \quad \text{if } \Delta x_1 \longrightarrow 0.\end{aligned} \quad (46)$$

So from (43), (46) can be

$$\left.\frac{\partial F_1}{\partial x_1}\right|_{(x_1,\dots,x_n)=(p_1,\dots,p_n)} = 0. \quad (47)$$

Then, taking (42) and (43) into (47), we have

$$\left.\frac{\partial F}{\partial x_1}\right|_{(x_1,\dots,x_n)=(p_1,\dots,p_n)} \ne 0. \quad (48)$$

Therefore, $(p_1,\dots,p_n)$ is not the extreme point of the function $F(x_1,\dots,x_n)$. For the same reason, $(q_1,\dots,q_n)$ is also not the extreme point.

Using the definition of extreme point, there exists a point $H = (h_1,\dots,h_n)$ such that $F(h_1,\dots,h_n) < F(p_1,\dots,p_n) = FD_{\alpha}(P,Q)$. As $F_1(h_1,\dots,h_n) = FD_{\alpha}(P,H)$, $F_2(h_1,\dots,h_n) = FD_{\alpha}(Q,H)$, then $FD_{\alpha}(P,H) + FD_{\alpha}(Q,H) < FD_{\alpha}(P,Q)$. The inequality is not consistent with (36).

From what has been discussed above, the conclusion in the theorem is obtained. □

5. Comparison between $FD_{1/2}$ and d

As [5, 6] mentioned, we focus on the RNA alphabet $\{U, C, A, G\}$. Code U as $(1,0,0,0)$: 1 shows that the first letter U is present, 0 shows that the second letter C does not appear, 0 shows that the third letter A does not appear, and 0 shows that the fourth letter G does not appear. Thereby, C is represented as $(0,1,0,0)$, A is represented as $(0,0,1,0)$, G is represented as $(0,0,0,1)$. So any codon can correspond to a fuzzy set as a point in the 12-dimensional fuzzy polynucleotide space I^{12}. For example, the codon CGU would be recorded as

$$(0,1,0,0,0,0,0,1,1,0,0,0) \in I^{12}. \quad (49)$$

However, there exist some cases in which there is no sufficient knowledge about the chemical structure of a particular sequence. One therefore may deal with base sequences not necessarily at a corner of the hypercube, and some components of the fuzzy set are not either 0 or 1. For example,

$$(0.3, 0.4, 0.1, 0.2, 0, 1, 0, 0, 0, 0, 0, 1) \in I^{12} \quad (50)$$

expresses a codon XCG. In this case, the first letter X is unknown and corresponds to U to an extent of 0.3, C to an extent of 0.4, A to an extent of 0.1, G to an extent of 0.2.

(1) For the metric d (in [5]),

d(histidine, proline) = d(CAU, CCG) = 0.8;

d(histidine, serine) = d(CAU, UCG) = 1;

d(histidine, arginine) = d(CAU, CGU) = 0.5;

d(arginine, glutamine) = d(CGU, CAG) = 0.8;

d(lysine, glycine) = d(AAA, GGG) = 1.

(2) For the metric $FD_{1/2}$,

$FD_{1/2}$(histidine, proline) = $FD_{1/2}$(CAU, CCG) = 1.6651;

$FD_{1/2}$(histidine, serine) = $FD_{1/2}$(CAU, UCG) = 2.0393;

$FD_{1/2}$(histidine, arginine) = $FD_{1/2}$(CAU, CGU) = 1.1774;

$FD_{1/2}$(arginine, glutamine) = $FD_{1/2}$(CGU, CAG) = 1.6651;

$FD_{1/2}$(lysine, glycine) = $FD_{1/2}$(AAA, GGG) = 2.0393.

From the above, the $FD_{1/2}$ is larger than d. But the value does not change the relationship of the distances between different codons. This shows that the new metric $FD_{1/2}$ reflects more information of the difference between codons. Next, the distances between codon XCG mentioned and proline and serine are as follows:

(1) For the metric d (in [5]),

d(XCG, proline) = d(XCG, CCG) = 0.3333;

d(XCG, serine) = d(XCG, UCG) = 0.3784;

(2) For the metric $FD_{1/2}$,

$FD_{1/2}$(XCG, proline) = $FD_{1/2}$(XCG, CCG) = 0.7408;

$FD_{1/2}$(XCG, serine) = $FD_{1/2}$(XCG, UCG) = 0.8271.

We apply the comparison to complete genomes. In [5], Torres a.nd Nieto computed the frequencies of the nucleotides A, C, G, and T at the three base sites of a codon in two bacteria *M. tuberculosis* and *E. coli* and obtained two points corresponding to either

$$(0.1632,\ 0.3089,\ 0.1724,\ 0.3556,\ 0.2036,\ 0.3145,\ 0.1763, 0.3056, 0.1645, 0.3461, 0.1593, 0.3302) \in I^{12}$$

or

$$(0.1605,\ 0.2420,\ 0.2600,\ 0.3374,\ 0.3116,\ 0.2286,\ 0.2846, 0.1752, 0.2619, 0.2568, 0.1831, 0.2981) \in I^{12}.$$

For the metric d (in [5]),

$$d\,(M.\ tuberculosis, E.coli) = 0.2483. \tag{51}$$

For the metric $FD_{1/2}$,

$$FD_{1/2}\,(M.\ tuberculosis, E.coli) = 0.2859. \tag{52}$$

It is easy to obtain that FD_α is closer to 1 and also larger than d when $0 < \alpha < 1/2$.

6. Concluding Remarks

By the discussion in the above sections, we come to the main conclusion: when P, Q are two points in the n-dimensional unit hypercube I^n, $FD_\alpha(P, Q)$ is a metric if and only if $0 < \alpha \leq 1/2$.

In Section 4, the method in case (iii) can also be used to prove that triangle inequality (36) does not hold in case (ii). But the method in case (ii) is intuitive, and we can find one determinate point R beyond the existence. So we adopt the method in case (ii) when $\alpha \geq 1$.

In this paper, we extend the method in [2, 4–6] to discuss the new fuzzy polynucleotide space. By considering all the possible values of parameter α, we obtain the new class of metrics in the space. At last, we numerically compare the new metrics FD_α to the former metric d by computing some basic examples of codons. This shows the improvement is comprehensive.

With $0 < \alpha \leq 1/2$, we can also study the metric space (I^n, FD_α) using the theory of metric space, such as the Pythagoras theorem, the isometric property, the isomorphism property, and the limit property in the future. We think the new metrics can interpret more biological significance for the sequences of the polynucleotide and be useful in the bioinformatics.

Acknowledgment

This paper is supported by the Fundamental Research Funds for the Central Universities (FRF-CU) Grant no. 2722013JC082.

References

[1] L. A. Zadeh, "Towards a theory of fuzzy systems," in *Aspects of Net-Works and Systems Theory*, R. E. Kalman and R. N. DeClairis, Eds., pp. 469–490, Holt, Rinehart & Winston, New York, NY, USA, 1971.

[2] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice Hall, Englewood Cliffs, NJ, USA, 1992.

[3] B. Kosko, *Fuzzy Engineering*, Prentice Hall, Englewood Cliffs, NJ, USA, 1997.

[4] K. Sadegh-Zadeh, "Fuzzy genomes," *Artificial Intelligence in Medicine*, vol. 18, no. 1, pp. 1–28, 2000.

[5] A. Torres and J. J. Nieto, "The fuzzy polynucleotide space: basic properties," *Bioinformatics*, vol. 19, no. 5, pp. 587–592, 2003.

[6] J. J. Nieto, A. Torres, and M. M. Vázquez-Trasande, "A metric space to study differences between polynucleotides," *Applied Mathematics Letters*, vol. 16, no. 8, pp. 1289–1294, 2003.

[7] A. Dress, T. Lokot, and L. D. Pustyl'nikov, "A new scale-invariant geometry on L_1 spaces," *Applied Mathematics Letters*, vol. 17, no. 7, pp. 815–820, 2004.

[8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley-Interscience, New York, NY, USA, 2nd edition, 2006.

[9] U. Kumar, V. Kumar, and J. N. Kapur, "Some normalized measures of directed divergence," *International Journal of General Systems*, vol. 13, no. 1, pp. 5–16, 1986.

[10] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Transactions on Information Theory*, vol. 37, no. 1, pp. 145–151, 1991.

[11] M. C. Pardo and I. Vajda, "About distances of discrete distributions satisfying the data processing theorem of information theory," *IEEE Transactions on Information Theory*, vol. 43, no. 4, pp. 1288–1293, 1997.

[12] F. Österreicher and I. Vajda, "A new class of metric divergences on probability spaces and its applicability in statistics," *Annals of the Institute of Statistical Mathematics*, vol. 55, no. 3, pp. 639–653, 2003.

[13] D. M. Endres and J. E. Schindelin, "A new metric for probability distributions," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1858–1860, 2003.

[14] S. W. Ho and R. W. Yeung, "On the discontinuity of the Shannon information measures," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5362–5374, 2009.

[15] F. Nielsen and S. Boltz, "The Burbea-Rao and Bhattacharyya centroids," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5455–5466, 2011.