Construction of a pre-contrast function on a deformed exponential family

Hiroshi Matsuzoe
Nagoya Institute of Technology
Department of Computer Science and Engineering
Gokiso-cho, Showa-ku, Nagoya, 466-8555, Japan
e-mail: *matsuzoe@nitech.ac.jp*

Abstract—A pre-contrast function is a direction dependent distance like function. Such a function naturally arises in the theory of non-conservative statistical inference or in quantum information geometry. On the other hand, a deformed exponential family is a set of probability distributions which contains long tailed probability distributions. Such probability distributions play important roles in the study of complex systems or in anomalous statistical physics. For a long tailed probability distribution, the standard expectation does not exist in general. Hence the notion of escort expectation has been introduced.

In this paper, we apply an escort expectations to pre-contrast functions. That is, we construct a pre-contrast function on a deformed exponential family by means of an escort expectation.

After giving preliminaries of geometry of contrast functions and pre-contrast functions, we review foundations of anomalous statistics. In particular, we study the notion of escort expectations. Then we construct a pre-contrast function on a set of non-exponential type probability density functions.

Keywords—information geometry; statistical manifold; contrast function; pre-contrast function; deformed exponential family; escort expectation

I. INTRODUCTION

Information geometry is a differential geometrical method for statistical inferences. A set of parametric probability density functions is called a statistical model, and it is regarded as a Riemannian manifold. However, the Levi-Civita connection with respect to the above Riemannian structure may not be useful for information gometry. A pair of dual affine connections are introduced and estimating procedures are clearly described using these dual affine connections [1], [2]. A contrast function is an asymmetric squared distance like function, and this asymmetry induces such a dualistic geometric structure of affine connections [4].

In the quantum version of information geometry or in advanced studies of statistical inferences, contrast functions may not exist in general. Therefore the notion of pre-contrast function was introduced to describe those geometric structures [5].

On the other hand, studies of non-standard phenomenon are important in recent statistical sciences. Anomalous statistics is one of research area which studies such non-standard phenomenon [14], [15]. In anomalous statistics, long tailed probability distributions and their related distributions are important objects. A deformed exponential family is a statistical model which includes such non-standard probability distributions [10], [11], [13].

In this paper, after reviewing geometry of contrast functions and pre-contrast functions, we focus on geometry of anomalous statistics. Then we construct a pre-contrast function for a deformed exponential family. A deformed expectation, called an escort expectation, plays important role in this framework.

II. STATISTICAL MANIFOLDS

We assume that all the objects are smooth throughout this paper. We start reviewing the definition of statistical manifolds.

Let M be a manifold of dimension n. Let h be a semi-Riemannian metric on M, and ∇ be an affine connection on M. Denote by T the torsion tensor field of ∇.

Definition II.1 (statistical manifold)**.** For given geometric objects (M, ∇, h), we assume that the following relation holds.

$$(\nabla_X h)(Y, Z) - (\nabla_Y h)(X, Z) = -h(T(X, Y), Z), \quad (1)$$

where X, Y and Z are arbitrary vector fields on M. In this case, we call the triplet (M, ∇, h) a *statistical manifold admitting torsion*. (We sometimes use an abbreviation "SMAT" for simplicity.) If the torsion tensor field vanishes everywhere on M, we say that the triplet (M, ∇, h) is a *statistical manifold*.

As we will see latter in this paper, a statistical manifold structure naturally arises in geometry of statistical model. Historically, the notion of statistical manifold was originally introduced by Lauritzen [7]. He called the triplet (M, g, C) a statistical manifold, where (M, g) is a Riemannian manifold and C is a totally symmetric $(0, 3)$-tensor field. Subsequently, Kurose [6] redefined a statistical manifold from the viewpoint of affine differential geometry.

For a statistical manifold admitting torsion (M, ∇, h), we define the *dual connection* ∇^* of ∇ with respect to h by

$$Xh(Y, Z) = h(\nabla^*_X Y, Z) + h(Y, \nabla_X Y). \quad (2)$$

Since h is nondegenerate, the dual connection ∇^* can be determined uniquely. The dual of the dual connection is the original one, that is, $(\nabla^*)^* = \nabla$, since h is symmetric.

This research was partially supported by JSPS (Japan Society for the Promotion of Science), KAKENHI (Grants-in-Aid for Scientific Research) Grant Numbers JP26108003 and JP15K04842

Moreover, the dual connection ∇^* is always torsion-free. In fact, from the definition of the dual connection (2), we have

$$\begin{aligned} & (\nabla_X h)(Y,Z) - (\nabla_Y h)(X,Z) \\ = \; & Xh(Y,Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z) \\ & -Yh(X,Z) + h(\nabla_Y X, Z) + h(Y, \nabla_Z X) \\ = \; & h(\nabla^*_X Y, Z) - h(\nabla_X Y, Z) \\ & -h(\nabla^*_Y X, Z) + h(\nabla_Y X, Z) \\ = \; & h(T^*(X,Y),Z) - h(T(X,Y),Z), \end{aligned}$$

where T^* is the torsion tensor filed of ∇^*. From Equation (1), T^* vanishes everywhere on M, and this implies that ∇^* is torsion-free.

We suppose that R and R^* are the curvature tensors of ∇ and ∇^*, respectively. From Equation (2), we have

$$h(R(X,Y)Z,V) + h(Z, R^*(X,Y)V) = 0.$$

Therefore, ∇ is curvature-free if and only if ∇^* is curvature-free. We say that a statistical manifold admitting torsion (M, ∇, h) is a *space of distant parallelism* if ∇ is curvature-free. That is, $R = R^* = 0$, $T^* = 0$, but $T \neq 0$.

Let us review examples of statistical manifolds [8].

Example II.1 (Riemannian manifold with the Levi-Civita connection)**.** Let (M, g) be a Riemannian manifold, and $\nabla^{(0)}$ be the Levi-Civita connection with respect to g. Then the triplet $(M, g, \nabla^{(0)})$ is a *trivial* statistical manifold since $\nabla^{(0)}$ is torsion-free and it is a metric connection. The dual connection ∇^* of $\nabla^{(0)}$ coincides with the original connection $\nabla^{(0)}$, that is, $\nabla^* = \nabla^{(0)}$.

Example II.2 (statistical model)**.** Let Ω be a total sample space. Suppose that $p(x;\xi)$ is a positive probability density distribution on Ω with parameter $\xi = (\xi^1, \ldots, \xi^n) \in \Xi \subset \mathbb{R}^n$, where Ξ is an open subset in $\mathbb{R}^n$. The set of all such probability densities S is called a *statistical model*, that is,

$$S = \left\{ p(x;\xi) \,\middle|\, \int_\Omega p(x;\xi)dx = 1,\ p(x;\xi) > 0,\ \xi \in \Xi \right\}. \quad (3)$$

We regard S is a manifold, and (ξ^i) is its local coordinate system. (See [2] for further details.)

Let $f(x)$ be a function on Ω. For $p \in S$, we denote by $E_p[f]$ the expectation of $f(x)$ with respect to $p(x;\xi)$. By setting $\partial_i = \partial/\partial\xi^i$, $l_\xi = \log p(x;\xi)$, we define a Riemannian metric $g^F = (g^F_{ij})$ on S, called a *Fisher metric*, by

$$g^F_{ij}(\xi) \quad = \quad E_p[\partial_i l_\xi \partial_j l_\xi].$$

We remark that $\partial_i l_\xi$ is called the *score function* of $p(x;\xi)$. Intuitively, it is a tangent vector of statistical model S at point $p(x;\xi)$. (cf. [5], [10], [11])

For a fixed $\alpha \in \mathbb{R}$, we define a torsion-free affine connection $\nabla^{(\alpha)}$ on S, called the *α-connection*, by

$$\Gamma^{(\alpha)}_{ij,k}(\xi) = E_p\left[\left(\partial_i\partial_j l_\xi + \frac{1-\alpha}{2}\partial_i l_\xi \partial_j l_\xi \right)(\partial_k l_\xi) \right],$$

where $\{\Gamma^{(\alpha)}_{ij,k}(\xi)\}$ are the Christoffel symbols of the first kind for $\nabla^{(\alpha)}$. We can check that $\nabla^{(0)}$ is the Levi-Civita connection with respect to the Fisher metric g^F and that $\nabla^{(\alpha)} g^F$ is totally symmetric. Hence the triplet $(S, \nabla^{(\alpha)}, g^F)$ is a statistical manifold. We call $(S, \nabla^{(\alpha)}, g^F)$ an *invariant statistical manifold* of S. It is known that $(S, \nabla^{(\alpha)}, g^F)$ is independent of the choice of reference measure on Ω [1], [2].

Example II.3 (quantum state space with SLD Fisher metric)**.** Let $\mathrm{Herm}(n)$ be the set of all Hermitian matrices of degree n. Suppose that S is the set of positive definite Hermitian matrices whose traces equal one, that is,

$$S = \{ P \in \mathrm{Herm}(n) \mid P > 0,\ \mathrm{tr}P = 1 \}.$$

Intuitively, S is regarded as the set of finite dimensional quantum states. Let T_0 be the set of traceless Hermitian matrices T_0:

$$T_0 = \{ X \in \mathrm{Herm}(n) \mid \mathrm{tr}X = 0 \}.$$

For an arbitrary point P in S, we identify T_0 the tangent space $T_P S$. We denote by $\mathcal{X}$ the vector field corresponding to the tangent vector $X \in T_0$.

For an arbitrary point P in S and an arbitrary tangent vector X in T_0, we define a *symmetric logarithmic derivative* $\omega_P(\mathcal{X})$ in $\mathrm{Herm}(n)$ by

$$X = \frac{1}{2}(P\omega_P(\mathcal{X}) + \omega_P(\mathcal{X})P).$$

The symmetric logarithmic derivative $\omega_P(\mathcal{X})$ is an analogy of a score function $\partial_i \log p(x;\xi)$ in classical statistical model.

A Riemannian metric g, called a *SLD Fisher metric*, is defined by

$$g_P(\mathcal{X},\mathcal{Y}) \quad = \quad \frac{1}{2}\mathrm{tr}\left(P(\omega_P(\mathcal{X})\omega_P(\mathcal{Y}) + \omega_P(\mathcal{Y})\omega_P(\mathcal{X})) \right).$$

An affine connection ∇ is defined by

$$(\nabla_{\mathcal{X}}\mathcal{Y})_P \quad = \quad h_P(\mathcal{X},\mathcal{Y})P - \frac{1}{2}(X\omega_P(\mathcal{Y}) + \omega_P(\mathcal{Y})X).$$

The connection ∇ is curvature-free but it is not torsion-free in general. This implies that the triplet (S, ∇, g) is a statistical manifold admitting torsion, and it is a space of distant parallelism.

III. Contrast functions and pre-contrast functions

A contrast function is an asymmetric squared distance like function, and a pre-contrast function is a direction depend distance like function. In this section, we review geometry of contrast and pre-contrast functions.

Let D be a function on $M \times M$. For arbitrary points p and q in M, and arbitrary vector fields $X_1, \ldots X_i$ and $Y_1, \ldots Y_j$ on M, we define a function $D[X_1 \cdots X_i | Y_1 \cdots Y_j]$ on M by the following formula.

$$\begin{aligned} & D[X_1 \cdots X_i | Y_1 \cdots Y_j](r) \\ & := (X_1)_p \cdots (X_i)_p (Y_1)_q \cdots (Y_j)_q D(p,q)|_{\substack{p=r \\ q=r}}. \end{aligned}$$

That is, $(X_1)_p \cdots (X_i)_p$ differentiate the first argument, and $(Y_1)_q \cdots (Y_j)$ differentiate the second argument on $M \times M$, then consider the restriction onto the diagonal $\{(r,r) | r \in M\}$. We say that D is a *contrast function* on M if it satisfies the following conditions.

1) For an arbitrary point $p \in M$, $D(p,p) = 0$.
2) $D[X|] = D[|X] = 0$.
3) $h(X,Y) := -D[X|Y]$ is a semi-Riemannian metric on M.

These conditions are sometimes called the *Eguchi relation* [4]. We say that g is the *induced Riemannian metric* on M from a contrast function D.

Proposition III.1. *For a contrast function D on M, we define ∇ and ∇^* by*

$$\begin{aligned} h(\nabla_X Y, Z) &:= -D[XY|Z], \\ h(Y, \nabla^*_X Z) &:= -D[Y|XZ]. \end{aligned}$$

Then ∇ and ∇^* are torsion-free affine connections on M mutually dual with respect to the induced metric h. Moreover, ∇h and ∇h are totally symmetric. Therefore, the triplets (M, ∇, h) and (M, ∇^*, h) are statistical manifolds.

Next, let us define a pre-contrast function. Let ρ be a function on $TM \times M$. A similar argument of contrast function, we define a function $\rho[X_1 \cdots X_i Z | Y_1 \cdots Y_j]$ on M by the following formula.

$$\begin{aligned} &\rho[X_1 \cdots X_i Z | Y_1 \cdots Y_j](r) \\ &:= (X_1)_p \cdots (X_i)_p (Y_1)_q \cdots (Y_j)_q \rho(Z_p, q)|_{\substack{p=r \\ q=r}}, \end{aligned}$$

where p and q are arbitrary points in M, and $X_1, \ldots X_i$, $Y_1, \ldots Y_j$ and Z are arbitrary vector fields on M.

We say that ρ is a *pre-contrast function* on M if it satisfies the following conditions.

1) For arbitrary functions $f_1, f_2 \in C^\infty(M)$, $\rho(f_1 X_1 + f_2 X_2, q) = f_1 \rho(X_1, q) + f_2 \rho(X_2, q)$.
2) $\rho[X|] = 0$.
3) $h(X,Y) := -\rho[X|Y]$ is a semi-Riemannian metric on M.

If $D(p,q)$ is a contrast function on M, $X_p D(p,q)$ is a *trivial* pre-contrast function.

Proposition III.2. *For a pre-contrast function ρ on M, we can define affine connections ∇ and ∇^* by*

$$\begin{aligned} h(\nabla_X Y, Z) &:= -\rho[XY|Z], \\ h(Y, \nabla^*_X Z) &:= -\rho[Y|XZ]. \end{aligned}$$

Then ∇ and ∇^* are mutually dual with respect to h. Moreover, the dual connection ∇^* is torsion-free. Therefore, the triplet (M, ∇, h) is a statistical manifold admitting torsion.

Example III.1 (Euclidean distance)**.** Let $\mathbf{R}^n$ be the standard Euclidean space. Suppose that g^E is the standard inner product on $\mathbf{R}^n$, and that ∇^E is the standard flat affine connection on $\mathbf{R}^n$. Then the triplet $(\mathbf{R}^n, \nabla^E, g^E)$ is a statistical manifold. In this case, a half of the Euclidean distance

$$D(x,y) = \frac{1}{2}||x-y||^2$$

is a contrast function on $\mathbf{R}^n$ which induces the statistical manifold $(\mathbf{R}^n, \nabla^E, g^E)$. Its directional differential

$$\rho(\partial/\partial x^i, y) = (\partial/\partial x^i)_x D(x,y) = x^i - y^i$$

is a pre-contrast function on M.

Example III.2 (Kullback-Leibler divergence)**.** Let S be a statistical model on Ω defined by (3). For two probability densities $p_1 = p(x;\xi_1)$ and $p_2 = p(x;\xi_2)$ in S, the *Kullback-Leibler divergence* (or the *relative entropy*) on S is defined by

$$D_{KL}(p_1, p_2) := \int_\Omega p(x;\xi_2) \log \frac{p(x;\xi_2)}{p(x;\xi_1)} dx.$$

Then D_{KL} is a contrast function on S which induces dual statistical manifolds $(S, \nabla^{(1)}, g^F)$ and $(S, \nabla^{(-1)}, g^F)$. Its directional differential

$$\begin{aligned} &D_{KL}\left(\left(\partial/\partial \xi^i\right)_{p_1}, p_2 \right) &(4) \\ &= -\int_\Omega \left(\frac{\partial}{\partial \xi^i} \log p(x;\xi_1) \right) p(x;\xi_2) dx \\ &= -E_{\xi_2}\left[\frac{\partial}{\partial \xi^i} \log p(x;\xi_1) \right] &(5) \end{aligned}$$

is a pre-contrast function on S.

We remark that this pre-contrast function ρ_{KL} is important in the theory of statistical inference. Suppose that $X_1, \ldots, X_N$ are random variables independently identically distributed to $p(x;\xi)$. Then the equation

$$\frac{1}{N} \sum_{a=1}^N \frac{\partial}{\partial \xi^i} \log p(X_a;\xi) = 0 \qquad (6)$$

is called an *estimating equation*. We also remark that a non-trivial pre-contrast function and its use for statistical inference have been studied in [5].

IV. Pre-contrast functions and estimating functions for deformed exponential families

In this section, we construct a pre-contrast function on a statistical model under deformed expectations of random variables.

Let χ be a strictly increasing function from $\mathbf{R}_{++}$ to $\mathbf{R}_{++}$. We consider the following non-linear differential equation.

$$\frac{d}{dt} \exp_\chi t := \chi(\exp_\chi t).$$

The eigenfunction $\exp_\chi t$ is called a χ*-exponential function* (or a *deformed exponential function*) [13], the function χ is called the *deformation function* of χ-exponential. If χ is the identity function, the standard exponential function is recovered.

The inverse of a χ-exponential function is called a χ*-logarithm function* (or a *deformed logarithm function*), and it is defined by

$$\ln_\chi s := \int_1^s \frac{1}{\chi(t)} dt.$$

If χ is the identity function, obviously the standard logarithm function is recovered.

We define a statistical model S_χ by the following formula.

$$S_\chi := \left\{ p(x;\theta) \;\middle|\; p(x;\theta) = \exp_\chi \left[\sum_{i=1}^n \theta^i F_i(x) - \psi(\theta) \right] \right\},$$

where $F_1(x), \ldots, F_n(x)$ are functions on the sample space Ω, $\theta = {}^t(\theta^1, \ldots, \theta^n) \in \Theta \subset \mathbf{R}^n$ is a parameter, and $\psi(\theta)$ is the normalization of probability density function defined by $\int_\Omega p(x;\theta)dx = 1$.

In anomalous statistics, the standard expectation may not work effectively [3], [13]. We consider deformations of expectations.

Let $S_\chi = \{p(x;\theta)\}$ be a χ-exponential family. We define the *escort distribution* $P_\chi(x;\theta)$ of $p(x;\theta) \in S_\chi$ by

$$P_\chi(x;\theta) := \chi(p(x;\theta)),$$

and the *escort expectation* $E_{\chi,p}[f(x)]$ by

$$\begin{aligned} E_{\chi,p}[f(x)] &= \int_\Omega f(x)P_\chi(x;\theta)dx \\ &= \int_\Omega f(x)\chi(p(x;\theta))dx. \end{aligned}$$

We remark that the escort distribution $P_\chi(x;\theta)$ is not a probability distribution, but a positive valued distribution in general. However, it is useful in the study of anomalous statistics [9], [12].

Definition IV.1 (estimating function)**.** Suppose that $S = \{p(x;\xi)\,|\,\xi \in \Xi\}$ is a statistical model, and that $u(x;\xi)$ is a $\mathbf{R}^n$-valued function on $\Omega \times \Xi$. We say that $u(x;\xi)$ is an *(unbiased) estimating function* on S if, for arbitrary point $p \in S$, the following conditions are satisfied.

1) $E_{\chi,p}[u(x;\xi)] = 0$,
2) $E_{\chi,p}[||u(x;\xi)||^2] < \infty$,
3) $\det\left(E_{\chi,p}[\partial/\partial\xi u(x;\xi)]\right) \neq 0$.

The first condition implies that the estimating function $u(x;\xi)$ is unbiased with respect to the escort expectation. The third condition is the distinguishability of estimating functions. In other ward, functions

$$\left\{\frac{\partial u}{\partial \xi^1}(x;\xi),\,\ldots,\,\frac{\partial u}{\partial \xi^n}(x;\xi)\right\}$$

should be linearly independent as functions on Ω.

Theorem IV.1. *Let $S_\chi = \{p(x;\theta)\}$ be a χ-exponential family, and $u(x;\theta)$ be an unbiased estimating function. Then, for two points $p = p(x;\theta_1)$ and $q = p(x;\theta_2)$ in S_χ,*

$$\begin{aligned} \rho_u((\partial/\partial\theta_1^i)_p, q) &:= -\int_\Omega u^i(x;\theta_1)\chi\{p(x;\theta_2)\}dx \quad (7) \\ &= -E_{\chi,q}\left[u^i(x;\theta_1)\right] \end{aligned}$$

is a pre-contrast function on S_χ.

Proof: 1). From fundamental properties of integration, we obtain $\rho_u(f_1X_1 + f_2X_2, q) = f_1\rho_u(X_1, q) + f_2\rho_u(X_2, q)$.

2). Since $u(x;\theta)$ is an unbiased estimating function, we have $\rho[X|] = 0$.

3) By differentiating (7), since $\det\left(E_{\chi,p}[\partial/\partial\xi u(x;\xi)]\right) \neq 0$. we obtain that $h(X,Y) = -\rho[X|Y]$ is a semi-Riemannian metric on S_χ. ■

At the end of this paper, we give an example of contrast function on a deformed exponential family.

Example IV.1 (χ-relative entropy)**.** Let S_χ be a deformed exponential family. For two probability densities $p_1 = p(x;\theta_1)$ and $p_2 = p(x;\theta_2)$ in S_χ, we define the χ*-relative entropy* by

$$\begin{aligned} D^\chi(p_1,p_2) &:= \int_\Omega \chi(p_2)[\ln_\chi p_2 - \ln_\chi p_1]dx \\ &= E_{\chi,p_2}[\ln_\chi p(x;\theta_2) - \ln_\chi p(x;\theta_1)]. \end{aligned}$$

If the deformation function χ is identity D^χ is nothing but a Kullback-Leibler divergence D_{KL}. If χ is a power function $\chi(t) = t^\beta$, D^χ is as the *α-divergence* in information geometry [1], [2].

By differentiating the χ-relative entropy, we have

$$\rho^\chi\left(\left(\partial/\partial\theta^i\right)_{p_1}, p_2\right) = -E_{\chi,p_2}\left[\frac{\partial}{\partial\theta^i}\ln_\chi p(x;\xi_1)\right]. \quad (8)$$

Then ρ^χ is a pre-contrast function on S_χ.

We remark again that an escort expectation plays an important role on a deformed exponential family. If we consider the standard expectation in Equation (8), $(\partial/\partial\theta^i)\ln_\chi p(x;\xi_1)$ is biased, so the function ρ^χ may not be a pre-contrast function.

References

[1] S. Amari, *Information Geometry and Its Applications*; Springer; Tokyo, 2016.
[2] S. Amari and H. Nagaoka, *Method of Information Geometry*, Amer. Math. Soc., Oxford University Press, Oxford, 2000
[3] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems: An Introduction*; Cambridge University Press, Cambridge, UK, 1993.
[4] S. Eguchi, *Geometry of minimum contrast*, Hiroshima Math. J., **22**(1992), 631–647.
[5] M. Henmi and H. Matsuzoe, *Geometry of pre-contrast functions and non-conservative estimating functions*, AIP Conf. Proc. 1340: International Workshop on Complex Structures, Integrability and Vector Fields, Amer. Inst. Phys., **1340**(2011), 32 – 41.
[6] T. Kurose, *On the divergences of* 1*-conformally flat statistical manifolds*, Tôhoku Math. J., **46**(1994), 427 – 433.
[7] S. L. Lauritzen, *Statistical manifolds*, Differential Geometry in Statistical Inferences, IMS Lecture Notes Monograph Series **10**, Hayward, California, 1987, 96 – 163.
[8] H. Matsuzoe, *Statistical manifolds and affine differential geometry*, Adv. Stud. Pure Math., **57**(2010), 303-321.
[9] H. Matsuzoe, *A sequence of escort distributions and generalizations of expectations on q-exponential family*, *Entropy*, **19**, no. 1, 7, (2017).
[10] H. Matsuzoe and M. Henmi, *Hessian structures on deformed exponential families* Lecture Notes in Comp. Sci., **8085**(2013), 275 – 282.
[11] H. Matsuzoe and M. Henmi, *Hessian Structures and Divergence Functions on Deformed Exponential Families*, Geometric Theory of Information, Signals and Communication Technology, Springer, (2014), 57 – 80.
[12] H. Matsuzoe and T.Wada, *Deformed algebras and generalizations of independence on deformed exponential families* Entropy, **17**, no. 8, (2015), 5729 – 5751.
[13] J. Naudts, Estimators, escort probabilities, and ϕ-exponential families in statistical physics, *J. Inequal. Pure Appl. Math.*, **5**, no. 102, (2004).
[14] J. Naudts, *Generalised Thermostatistics*, Springer-Verlag, 2011.
[15] C. Tsallis, *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*, Springer, New York, 2009.
[16] T. Wada and H. Matsuzoe *Conjugate representations and characterizing escort expectations in information geometry*, Entropy, **19**, no. 7, 309, (2017).