

Article

Tensor Properties of Joint Probability Densities on Riemannian Parametric Manifold

Manouchehr Amiri

Health Organization, Tehran, Iran; manoamiri@gmail.com

Abstract: We show the tensor properties of joint probabilities on a Riemannian parametric manifold. Initially we develop a binary data matrix for parameter measurements of a large number of particles confining in a closed system in order to retrieve the joint probability densities of related parameters. By introducing a new generalized inner product as a multilinear operation on dual basis of parametric space, we extract the set of joint probabilities and prove them to meet contravariant tensor properties in a general Riemannian parametric space. We show these contravariant tensors reduce to classical ordinary partial derivatives definition in ordinary Euclidean parametric space. Finally we prove by a theorem that symmetrized iterative contravariant derivative of cumulative probability function on Riemannian manifold gives the set of joint probabilities in those manifold. We bring some examples for compatibility with physical tensors.

Keywords: joint probability density; tensor properties; probability on Riemannian manifold; covariant derivative; contravariant derivative

Mathematics Subject Classification 2010: 53A35, 60A99

1 Introduction

The current classical theory of probability distributions have been based on the presumptive Euclidean space of parameters and random variables. In this sense the joint probability densities being derived from sequential partial derivatives of corresponding cumulative distribution function [1,2]. The non-Euclidean manifolds endowed with Riemannian metrics in the context of probability theory, have been introduced by many authors for phase and parameter space [3,9], [4]. Moreover using differential geometry, some variations of information theory having been devised on non-Euclidean (Riemannian) metric spaces in this field by Fisher, Rao, Amari and others [3,5]. In these approaches, probability distributions of various models exhibited as *points* on some Riemannian manifolds. By this background, differential geometry technique could be applied to analyze the probability distribution manifolds. The most applicable metric tensor on these spaces first introduced by Fisher and Rao the so-called Fisher information metric [6]. In this short article we introduce a new model based on binary data matrices to record the parameters of a system of particles. Each particle has its own properties and occupies just on infinitesimal interval

on each parameter coordinate x^ν . In sec (2) we develop the concept of *particle oriented coordinates* which span a flat Euclidean space on which we embed Riemannian parametric manifolds. Iteration of measurements for all particles and parameters yield a binary matrix containing all particle system information. In sec (3) we proceed by introducing a manifold whose base vectors are the rows of data matrix then we extract the joint probabilities of events based on the dual space base vectors and a new definition of *inner* products and prove their tensor properties. In sec (4) we prove the general joint probabilities on Riemannian parametric manifolds as the iterative contravariant derivatives of a scalar i.e. cumulative distribution function and proceed for higher order joint probabilities through symmetrized covariant derivatives. In sec (5) we bring some examples of the manifolds whose metrics developed by covariant derivative of a scalar.

Definition 1.1

Parametric space \mathcal{M} specified by x^ν coordinates with ν from 1 to the dimension of parametric space, d :

$$1 \leq \nu \leq d \quad (1.1)$$

So the number of involved independent parameters (degree of freedom) is d . The parametric space \mathcal{M} in present article, generally presumed to be a Riemannian manifold with local coordinates x^ν and basis vectors $e_\alpha = \frac{\partial}{\partial x^\alpha}$ at any point $\mathbf{p} \in \mathcal{M}$. If we divide the coordinate x^ν into m_ν small intervals $\Delta x^\nu(i)$, then integer (i) refers to the i th interval of this parameter. This means that i replaces the coordinate values of point \mathbf{p} and ranges between $\mathbf{0}$ and an integer m_ν :

$$1 \leq i \leq m_\nu \quad (1.2)$$

The whole parametric manifold includes all coordinates x^ν and associated basis vectors. The overall number of intervals reads as: $m = \sum_\nu m_\nu$ (1.3)

2 Introducing Particle oriented coordinate

Assume a system consisting of large number N of particles confined in an interval of space-time, undergoes measurements. Taking into account of such a system of particles, brings us the advantage of choosing unlimitedly a sufficient huge number of particles, otherwise we could substitute particles by any kind of systems defined by their arbitrary parametric space.

Suppose a set of related parameters labelled by ν is to be measured in an interval of time ΔT . We label the particles by integer numbers up to N , and divide the possible range of each parameter into such small intervals that satisfy the accuracy of the experiment. These intervals defined as in definition 1.1, denoted by $\Delta x^\nu(i)$ where i stands for the ordered location number of an interval on parameter coordinate x^ν :

$$x^\nu(i-1) \leq x^\nu(i) \leq x^\nu(i+1) \quad (2.1)$$

So the parameter value of each particle falls in just one of these intervals. For each interval $\Delta x^\nu(i)$, there are some particles that their parameter with label ν places in this interval. Let record the results of these outcomes in an array whose entries are $\mathbf{0}$ or $\mathbf{1}$ in such a way that for particles with

parameter value in interval $\Delta x^\nu(i)$, the corresponding value in array reads **1** and otherwise **0**. For instance for the first particle, if its value of parameter ν falls in the range of $\Delta x^\nu(i)$ we return **1**, and otherwise **0**. If this process iterates for all particles then we obtain an array of entries for this interval which could be arranged as a vector $e^{*\nu}(i)$. This vector is a binary array which carries the information of this interval for system of particles e.g.

$$e^{*\nu}(i) = (1,0,0,1,1,0,1,1, \dots) \quad (2.2)$$

Each of these **0** and **1** connected to a specific particle in the system and the total number of entries equals N . one may attribute to any particle, an independent coordinate and consider a N -dimensional manifold spanned by these coordinate in such a way that the result of projection of $e^{*\nu}(i)$ on these coordinates turns out the array in equation (2.2).

Here any particle specifies an independent coordinate with two possible values **0** and **1**. These coordinates are *orthogonal*, because at the initial setting the parameter values of every particle (such as position and momentum etc) considered to be independent of all other particles. We call these set of coordinate as *particle oriented coordinate* that as a coordinate chart is homeomorphic to a subset of Euclidean flat space \mathbb{R}^N which span a manifold \mathbf{M} . Moreover parametric space \mathcal{M} of considered system includes all coordinates x^ν and their dual basis $e^{*\nu}$ where the latter span a dual tangential (cotangent) vector space $T_p^*\mathcal{M}$ defined on a point \mathbf{p} in parametric space \mathcal{M} i.e.

$$Span\{e^{*\nu}\} = T_p^*\mathcal{M} \subset \mathbf{M} \quad (2.3)$$

The set of $\{e^{*\nu}(i) | 0 \leq i \leq m_\nu\}$ for a fixed ν as dual basis vectors for parameter x^ν , span a cotangent space $T_p^{*\nu}\mathcal{M}$ that is a sub-space of $T_p^*\mathcal{M}$:

$$T_p^*\mathcal{M} = \cup_\nu T_p^{*\nu}\mathcal{M} \quad (2.4)$$

Manifold \mathcal{M} is not necessarily flat and as shown in next sections will be considered as Riemannian manifolds with non-zero curvatures.

Definition 2.1

Any interval $\Delta x^\nu(i)$ specifies a vector $e^{*\nu}(i)$ whose components merely accept binary values of **0** and **1** in Vector space of *particle oriented coordinate*. Therefor $e^{*\nu}(i)$ in this coordinates may be represented by an array like:

$$e^{*\nu}(i) = (1,0,0,1,1,0,1,1, \dots) \quad (2.5)$$

Definition 2.2

For a fixed ν , we define the matrix $\mathfrak{D}_{m_\nu \times N}$ with $e^{*\nu}(i)$ as i th row. Obviously each column of this matrix contains just one **1**, because as described before each column belongs to a particle which occupies just one value (interval) $\Delta x^\nu(i)$.

Lemma 2.2 The set of basis $\{e^{*\nu}(i)\}$ for fixed ν are mutually orthogonal.

Proof: any particle takes just one value and consequently one interval on x^ν coordinate. So as was shown in definition 2.2, columns of $\mathfrak{D}_{m_\nu \times N}$ carries just one entry 1 while other entries takes 0. If

the n th component of $e^{*\nu}(i)$ be denoted by $[e^{*\nu}(i)]_n$, then the inner product $\langle e^{*\nu}(i), e^{*\nu}(j) \rangle$ can be read as:

$$\langle e^{*\nu}(i), e^{*\nu}(j) \rangle = \sum_n [e^{*\nu}(i)]_n [e^{*\nu}(j)]_n$$

$[e^{*\nu}(i)]_n$ and $[e^{*\nu}(j)]_n$ do not take the $\mathbf{1}$ value simultaneously, then it is easy to conclude that inner product $\langle e^{*\nu}(i), e^{*\nu}(j) \rangle$ vanishes for $i \neq j$ and orthogonality of these bases is proved.

Postulate 2.3

At the limit $\Delta x^\nu(i) \rightarrow dx^\nu(i)$, the vectors $e^{*\nu}(i)$ approaches the basis of cotangent bundle (space) as 1 -form i.e.

$$e^{*\nu}(i) \equiv dx^\nu(i) \equiv \omega^\nu(i) \quad (2.6)$$

Therefor the notation (2.6) exhibits the bases of a cotangent bundle (space) $T_p^{*\nu}M$ which is the dual of parametric space tangent bundle T_pM . As described before, these basis vectors specify the interval i on a specific parameter x^ν . The relative abundance of particles within this interval is proportional to the ratio of total number of $\mathbf{1}$ in the $e^{*\nu}(i)$, denoted by $n^\nu(i)$, to total particles number N :

$$p^\nu = \frac{n^\nu(i)}{N} \quad (2.7)$$

It is straight forward to conclude that $n^\nu(i)$ equals the inner product $\langle e^{*\nu}(i), e^{*\nu}(i) \rangle$. In an infinitesimal limit of $\Delta x^\nu(i)$ this dual basis vectors approaches to $dx^\nu(i)$ which belongs to the cotangent (dual) space basis of $T_p^{*\nu}M$ while carries the information about that interval in considered system. Respect to coordinate transformations $x^\alpha \rightarrow \bar{x}^\beta$ in parametric space we obtain the rule of $dx^\nu(i)$ transformation with summation on ν :

$$d\bar{x}^\mu(\bar{j}) = \frac{\partial \bar{x}^\mu(\bar{j})}{\partial x^\nu(i)} dx^\nu(i) \quad (2.8)$$

Or with the identity (2.6):

$$\bar{e}^{*\mu}(\bar{j}) = \frac{\partial \bar{x}^\mu(\bar{j})}{\partial x^\nu(i)} e^{*\nu}(i) \quad (2.9)$$

\bar{j} stands for the transformed locations of i in new coordinate \bar{x}^β . Since $dx^\nu(i)$ are considered infinitesimals, the number of particles confined to this interval always constitute a binary array even in transformed basis vectors $\bar{e}^{*\nu}$ and therefore these basis vectors in all coordinates contain a binary array similar to $e^{*\nu}(i)$.

3 Binary Data matrix:

Combining all the row vectors $e^{*\nu}(i)$ for all d parameters, gives a matrix $\mathbf{D}_{m \times N}$ with m rows and N columns where m defined in (1.3) and N stated as total particles number. This matrix itself is also the conjoint of d matrices each for a specific parameter. As an example if for a parameter with coordinate x^ν , there are m_ν intervals, then a matrix $\mathfrak{D}_{m_\nu \times N}$ with m_ν rows and N columns could be identified as a part of $\mathbf{D}_{m \times N}$ where $\{e^{*\nu}(i)\}$ (with fixed ν) constitute the set of dual basis. The columns of $\mathfrak{D}_{m_\nu \times N}$ carries just one entry $\mathbf{1}$, because each column corresponds to one particle that its parameter value falls in just one of intervals Δx^ν , namely $\Delta x^\nu(i)$. Obviously combining all $\mathfrak{D}_{m_\nu \times N}$ result in the data matrix $\mathbf{D}_{m \times N}$. The set of $\{e^{*\nu}(i)\}$ with fixed ν span a tangent space $T_M^{*\nu}$ on a point $p \in T_M^*$ (or \mathcal{M}).

Definition 3.1 In this model an *event* refers to placing the value of x^ν in the interval i of that parameter i.e. $x^\nu(i)$ for a particle.

Definition 3.2 Joint probability density of two events $x^\nu(i)$ and $x^\mu(j)$ is proportional to the number of particles that simultaneously have the same parameter values of $x^\nu(i)$ and $x^\mu(j)$ where i and j indicate the values that are taken by x^ν and x^μ parameters (i.e. coordinate values). If this number presented by $n^{\nu\mu}$ the exact form of joint probability density reads as:

$$f^{\nu\mu}(i, j) = \frac{n^{\nu\mu}}{N} \quad \nu \neq \mu \quad (3.1)$$

It is noteworthy to remind that for $\nu = \mu$ the joint density $f^{\nu\mu}$ and $n^{\nu\mu}$ reduce to f^ν and n^ν respectively.

Definition 3.3 We define the *generalized inner product* \odot for vectors U, V, W, \dots in a vector space \mathcal{V} with orthogonal local coordinates as a multilinear map:

$$\Phi: \mathcal{V} \rightarrow \mathcal{V}^* : U \odot V \odot W \odot \dots = \sum_n u_n v_n w_n \dots \quad (3.2)$$

This is simply a generalization of inner products in usual definition or a functional on dual vector space \mathcal{V} . consequently \mathcal{V}^* is dual of \mathcal{V} or double dual of coordinate space.

Lemma 3.4 Generalized scalar operation \odot is linear.

Due to the definition (3.3) it is straightforward to see that:

$$\begin{aligned} 1) & U \odot V = V \odot U \\ 2) & \lambda_1 U \odot \lambda_2 V = \lambda_1 \lambda_2 U \odot V \\ 3) & (U + W) \odot V = U \odot V + W \odot V \end{aligned} \quad (3.3)$$

It is noteworthy that this map lacks the *associative* property.

Lemma 3.5

Regarding definition (3.2) for generalized inner product of any subset of cotangent space basis $\{e^{*\mu}(i), e^{*\nu}(j), e^{*\xi}(k), \dots\}$, the joint probability of events $x^\mu(i)$, $x^\nu(j)$ and $x^\xi(k)$... can be given by equation:

$$f^{\mu\nu\xi\dots}(i, j, k, \dots) = \frac{1}{N} (e^{*\mu}(i) \odot e^{*\nu}(j) \odot e^{*\xi}(k) \odot \dots) \quad (3.4)$$

Proof: Using the equation (3.2) by omitting the location indices i, j, k, \dots we have:

$$e^{*\mu} \odot e^{*\nu} \odot e^{*\xi} \odot \dots = \sum_n (e^{*\mu})_n (e^{*\nu})_n (e^{*\xi})_n \dots \quad (3.5)$$

Summation carried out on all particles. Since components of $(e^{*\mu})_n$ take two values **0** or **1**; the non-vanishing terms are those with components that simultaneously take **1**, and therefore right side sum of equation (3.5) reduces to the number of particles whose parameter values simultaneously places in the intervals i, j, k, \dots on particles coordinate. Normalization of (3.4) by $\frac{1}{N}$ yields the ratio of this number to total particles number and consequently gives the joint probability of simultaneous occurring of events $x^\mu(i)$, $x^\nu(j)$ and $x^\xi(k)$.. For instance in our model inner products $e^{*\mu} \odot e^{*\nu} = \langle e^{*\mu}, e^{*\nu} \rangle$ stand for the number of particles which have the common parameter at i and j locations, or the joint probability of i and j in the related system of particles i.e. $n^{\nu\mu}$. For a general case we consider $n^{\nu\mu\xi}$ that represents the number of particles with

common parameters at locations i, j, k, \dots . Consequently $n^{\mu\nu\xi}$ with a normalization factor $\frac{1}{N}$ (N is the total number of particles) gives the joint probability of occurrence of i, j, k, \dots intervals in the above mentioned system:

$$f^{\mu\nu\xi..}(i, j, k, \dots) = \frac{n^{\mu\nu\xi}}{N} \quad (3.6)$$

This equation reveals a specific configuration and state of related system for which there is a specific set of joint probabilities $\{f^{\mu\nu\xi..}\}$ that represents the exact state of it. Any configuration (state) of this system represented by such a specific set and vice versa. Evidently the order of indices ν, μ, \dots does not effect on the related joint probability.

For a continuous density, $n^{\mu\nu\xi}$ is proportional to $f^{\mu\nu\xi..}$ and volume element dv and should be modified as:

$$f^{\mu\nu\xi..}(i, j, k, \dots)dv = \frac{n^{\mu\nu\xi}}{N} \quad (3.7)$$

As described in equation (3.3), we could attribute a quantity $f^{\mu\nu\xi..}$ to any multilinear form of basis vectors $e^{*\nu}(i)$ defined in (2.5), as follows:

$$f^{\mu\nu\xi} = \frac{1}{N} e^{*\mu} \odot e^{*\nu} \odot e^{*\xi} \odot \dots = \frac{1}{N} \sum_n (e^{*\mu})_n (e^{*\nu})_n (e^{*\xi})_n \dots \quad (3.8)$$

This holds true for any transformed basis $\bar{e}^{*\mu}$:

$$\bar{f}^{\mu\nu\xi..} = \frac{1}{N} \bar{e}^{*\mu} \odot \bar{e}^{*\nu} \odot \bar{e}^{*\xi} \odot \dots = \frac{1}{N} \sum_n (\bar{e}^{*\mu})_n (\bar{e}^{*\nu})_n (\bar{e}^{*\xi})_n \dots \quad (3.9)$$

Where $e^{*\mu}, e^{*\nu}, \dots$ considered at ordered i, j, k, \dots locations and $\bar{e}^{*\mu}, \bar{e}^{*\nu}, \dots$ at transformed ordered locations $\bar{i}, \bar{j}, \bar{k}, \dots$. The joint probabilities $f^{\mu\nu\xi}$ generally speaking, are invariant under the permutation of μ, ν, ξ, \dots indices, because this joint probabilities, axiomatically does not depend on the order of simultaneous events and their indices.

Lemma 3.6 $f^{\mu\nu\xi..}$ defined in (3.8) is a contravariant tensor.

Proof: with a coordinate transformation $x^\alpha \rightarrow \bar{x}^\beta$ we have:

$$\bar{e}^{*\mu} = \frac{\partial \bar{x}^\mu}{\partial x^\nu} e^{*\nu} \quad (3.10)$$

Substitution into (3.9) gives:

$$\bar{f}^{\mu\nu\xi..} = \frac{1}{N} \bar{e}^{*\mu} \odot \bar{e}^{*\nu} \odot \bar{e}^{*\xi} \odot \dots = \frac{1}{N} \sum_n \left(\frac{\partial \bar{x}^\mu}{\partial x^\alpha} e^{*\alpha} \right)_n \left(\frac{\partial \bar{x}^\nu}{\partial x^\beta} e^{*\beta} \right)_n \left(\frac{\partial \bar{x}^\xi}{\partial x^\gamma} e^{*\gamma} \right)_n \dots \quad (3.11)$$

Due to linear property of \odot , and the definition of $f^{\alpha\beta\gamma} = \frac{1}{N} e^{*\alpha} \odot e^{*\beta} \odot e^{*\gamma} \odot \dots$ we obtain:

$$\bar{f}^{\mu\nu\xi..} = \frac{1}{N} \left(\frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \frac{\partial \bar{x}^\xi}{\partial x^\gamma} \dots \right) e^{*\alpha} \odot e^{*\beta} \odot e^{*\gamma} \odot \dots = \left(\frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} \frac{\partial \bar{x}^\xi}{\partial x^\gamma} \dots \right) f^{\alpha\beta\gamma..} \quad (3.12)$$

This reveals the tensor property of $f^{\alpha\beta\gamma}$ while we use the Einstein summation convention. Actually $f^{\alpha\beta\gamma}(i, j, k, \dots)$ exhibits a large number of distributions for all permutation of i, j, k, \dots locations of parameters.

Corollary 3.7 while the indices i, j, k, \dots refer to the location numbers of intervals belong to some ordered parameters $x^\alpha(i)$, $x^\beta(j)$ and $x^\gamma(k)$, the rank of tensor $f^{\alpha\beta\gamma}$ equals the number of these indices. Then for any fixed $\alpha\beta\gamma \dots$ indices we have:

$$1 = \sum_{i,j,k,\dots} f^{\alpha\beta\gamma\dots}(i, j, k, \dots) \quad (3.13)$$

Proof: we have shown that tensors $f^{\alpha\beta\gamma\dots}(i, j, k, \dots)$ are the joint probabilities of $x^\mu(i)$, $x^\nu(j)$ and $x^\xi(j) \dots$ events. For these joint probabilities the equation (3.13) is valid as a general property [2,10].

Definition 3.8 Regarding the definition of cumulative distribution or joint distribution function (CDF) [2,10] we may define a function \mathcal{N} on any point $\mathbf{p}(x^1, x^2 \dots x^d) \in \mathcal{M}$ as follows:

- 1) $0 \leq \mathcal{N} \leq 1$
- 2) $\lim_{x^1, x^2 \dots x^d \rightarrow \infty} \mathcal{N} = 1$
- 3) $\lim_{x^1, x^2 \dots x^d \rightarrow -\infty} \mathcal{N} = 0$
- 4) $\mathcal{N} = \frac{N'}{N}$
- 5) Where N' stands for the total particle whose parameters take values below the $x^1, x^2 \dots x^d$

At the limit of $N \rightarrow \infty$ the function \mathcal{N} behaves as a smooth and differentiable function F on parametric space \mathcal{M} .

4 Joint probability as contravariant derivatives

It is usual in probability texts to explain the multiple variable joint probability density in Euclidean parameter space with Cartesian coordinate as a partial derivative of cumulative distribution function [1, 2,10]:

$$f^{\alpha\beta\gamma\dots} = \frac{\partial^n F}{\partial x_\alpha \partial x_\beta \partial x_\gamma \dots} = \partial^\alpha \partial^\beta \partial^\gamma \dots F \quad (4.1)$$

Lemma 4.1 Let F be a differentiable cumulative distribution function as defined in definition 3.8. In Taylor expansion of this function about a point (x_{01}, x_{02}, \dots) , regarding (4.1) in Cartesian coordinate, the probability densities $f^{\alpha\beta\gamma\dots}$ appear as the coefficients of powers $(x_\alpha - x_{0\alpha})(x_\beta - x_{0\beta})(x_\gamma - x_{0\gamma}) \dots$ while $(x_\alpha - x_{0\alpha})$ and $(x_\beta - x_{0\beta}) \dots \rightarrow 0$.

Eventually the equation (4.1) is working for a Cartesian coordinates of parameters in a Euclidean manifold while F is cumulative probability or joint distribution function [1, 2,10]. In such a space, the reciprocal dual coordinates x_α coincides the original coordinate x^α . Nevertheless equation (4.1) is not valid for coordinate of a Riemannian space. It is well known that this partial derivative generally is not a tensor except for its first order and in a Euclidean space with Cartesian coordinate. Naturally the first choice for substitution of equation (4.1) in Riemannian space will be fitted by covariant derivatives. However because the tensors $f^{\mu\nu\xi\dots}$ in lemma 3.6 are contravariant tensors, we propose a contraction of *covariant derivatives* with metric tensor, to derive *contravariant derivatives* for the general form of joint probability in parametric space with Riemannian manifold properties. Actually these contravariant derivatives of any order remain as tensors:

$$\nabla^i \nabla^j \nabla^k \dots F \quad (4.2)$$

here we use contraction of covariant derivative ∇_m with metric tensor i.e. $\nabla^i = g^{im} \nabla_m$ and call it as *contravariant derivative* ∇^i . Recalling that because of lower indices symmetry for torsion free connection of covariant derivative, the successive operation of covariant derivative will not generally commutative except for double covariant derivative of a scalar. Consequently for contravariant derivative we have the same commutativity:

Lemma 4.2 commutative property of contravariant derivative holds for second derivative of a scalar:

$$\nabla^i \nabla^j F = \nabla^j \nabla^i F \quad (4.3)$$

Proof: after using contraction with metric tensor we obtain in terms of covariant derivative:

$$\nabla^i \nabla^j F = g^{im} \nabla_m g^{jn} \nabla_n F$$

Because the covariant derivative of metric tensor vanishes, metric tensor acts as a constant with respect covariant derivative:

$$\nabla^i \nabla^j F = g^{im} g^{jn} \nabla_m \nabla_n F$$

Using the commutative property of covariant derivative we have:

$$\nabla^i \nabla^j F = g^{im} g^{jn} \nabla_m \nabla_n F = g^{im} g^{jn} \nabla_n \nabla_m F = \nabla^j \nabla^i F$$

Naturally it is supposed that the order of events does not matter for a joint probability of some events. Although the commutativity for two events guarantees it for the case of two events measurements, however for joint probabilities of more than two measurements (events), the commutativity of contravariant derivatives will not met. On the other hand regard to lemma (3.6) joint probability $f^{\mu\nu\xi}$ as a tensor is *full symmetric* respect to contravariant indices. Consequently the derivatives should be symmetrized to fulfil this property of joint probabilities. Symmetrization of multiple contravariant derivative (symmetrized covariant derivative) could be accomplished by routine symmetric form:

$$\nabla^{(i} \dots \nabla^{k)} \quad (4.4)$$

Which stands for the sum over all permutations of indices.

Corollary 4.2 The set of all paired joint probabilities constitute the metric tensor of dual space \mathcal{M}^* .

Proof: regarding equation (3.3) for paired joint probabilities we have:

$$f^{\mu\nu}(i, j) = \frac{1}{N} (e^{*\mu}(i) \odot e^{*\nu}(j)) \quad (4.5)$$

Right side stands for ordinary inner product of dual space basis vectors. Respect to definition of metric tensor, these inner products are proportional to metric tensor of related manifold:

$$f^{\mu\nu} \sim g^{*\mu\nu} = \langle e^{*\mu}, e^{*\nu} \rangle \quad (4.6)$$

Knowing the metric of dual space \mathcal{M}^* reveals all properties of manifold (including curvatures). Taking into account the identity $f^{\mu\nu} \sim g^{*\mu\nu}$ we conclude that $f^{\mu\nu}$ are sufficient to estimate the most part of information about other joint probabilities and distribution of related system. We use the notation g and g^* as determinants of metric tensors g_{ij} and g^{*ij} respectively.

Theorem 4.3 $f^{ijk..}$ in a general Riemannian manifold coordinates of parametric space could be represented as an iteration of a symmetrized *contravariant* derivative of a scalar i.e. cumulative probability distribution as a continuous variable which replaces F in (4.1). However we have shown that joint probability $f^{ijk..}$ is a tensor that in Euclidean space with Cartesian coordinates reduces to partial derivative of equation (4.1) and therefore for a general Riemannian manifold the best choice is a symmetrized contravariant derivative instead of partial derivatives of cumulative probability as follows:

$$f^{\mu_1\mu_1\dots\mu_k} = \nabla^{(\mu_1\mu_1\dots\mu_k)} F \quad (4.7)$$

Proof: Respect to Taylor covariant derivative expansion of a scalar on Riemannian manifold we have [11,12]:

$$\varphi = \sum_{k=1}^d \frac{1}{k!} (\tau - \tau_0)^k \left(u^{\mu_1} u^{\mu_2} \dots u^{\mu_k} \nabla_{(\mu_1} \dots \nabla_{\mu_k)} \varphi \right) \Big|_{\tau=\tau_0} \quad (4.8)$$

Where Symmetrization imposed on iterated covariant derivatives. With definition $u^\mu = \frac{dx^\mu}{d\tau}$ at an infinitesimal interval $\tau - \tau_0 = \Delta\tau \rightarrow 0$ the above equation reduces to:

$$\varphi = \sum_{k=1}^d \frac{1}{k!} \left(dx^{\mu_1} \dots dx^{\mu_k} \nabla_{(\mu_1} \dots \nabla_{\mu_k)} \varphi \right) \Big|_{\tau=\tau_0} \quad (4.9)$$

$$\varphi = \sum_{k=1}^d dx^{\mu_1} \dots dx^{\mu_k} \left(\frac{1}{k!} \nabla_{(\mu_1} \dots \nabla_{\mu_k)} \varphi \right) \Big|_{\tau=\tau_0} \quad (4.10)$$

Replacing φ with cumulative distribution function F and $(dx^{\mu_1} \dots dx^{\mu_k})$ with volume element $\frac{1}{\sqrt{g}} dv^k$ yields (g is metric tensor determinant):

$$dv^k = \sqrt{g} dx^{\mu_1} \dots dx^{\mu_k}$$

$$F = \frac{1}{\sqrt{g}} \sum_{k=1}^d dv^k \left(\frac{1}{k!} \nabla_{(\mu_1} \dots \nabla_{\mu_k)} F \right) \Big|_{\tau=\tau_0} \quad (4.11)$$

Where g is metric tensor determinant. The last term in this sum stands for the probability of being in volume element dv^k . Therefore respect to (3.6) and after replacing covariant derivatives in (4.11) by contravariant derivatives through contraction with metric tensor:

$$\nabla_\mu = g_{\mu\nu} \nabla^\nu \quad (4.12)$$

(4.11) converts to:

$$F = \frac{1}{\sqrt{g}} \sum_{k=1}^d dv^k (g_{\mu_1\nu_1} g_{\mu_2\nu_2} \dots) \left(\frac{1}{k!} \nabla^{(\nu_1} \dots \nabla^{\nu_k)} F \right) \Big|_{\tau=\tau_0} \quad (4.13)$$

We used the well-known property of covariant derivative: $\nabla_\mu g_{\nu\theta} = 0$ which makes the metric tensor in covariant derivative, as constants in ordinary derivatives. So $g_{\mu_1\nu_1}$ could be factorized out from a covariant derivative term.

With transforming dv^k to contravariant form:

$$dx^{\mu_1} \dots dx^{\mu_k} = (g^{\mu_1\nu_1} dx_{\nu_1}) (g^{\mu_2\nu_2} dx_{\nu_2}) \dots = (g^{\mu_1\nu_1} g^{\mu_2\nu_2} \dots) (dx_{\nu_1} dx_{\nu_2} \dots) \quad (4.14)$$

With definition $dv_k = \sqrt{g'} dx_{v_1} dx_{v_2} \dots$ (g' is dual metric determinant) and replacing into equation (4.13) gives:

$$F = \frac{1}{\sqrt{g'}} \sum_{k=1}^d dv_k (g^{\mu_1 \nu_1} g^{\mu_2 \nu_2} \dots) (g_{\mu_1 \nu_1} g_{\mu_2 \nu_2} \dots) \left(\frac{1}{k!} \nabla^{(\nu_1} \dots \nabla^{\nu_k)} F \right) \Big|_{\tau=\tau_0} \quad (4.15)$$

$$F = \frac{1}{\sqrt{g'}} \sum_{k=1}^d dv_k (g^{\mu_1 \nu_1} g_{\mu_1 \nu_1}) (g^{\mu_2 \nu_2} g_{\mu_2 \nu_2}) \dots \left(\frac{1}{k!} \nabla^{(\nu_1} \dots \nabla^{\nu_k)} F \right) \Big|_{\tau=\tau_0} \quad (4.16)$$

By performing Einstein Summation convention on indices μ_i, ν_i the terms $g^{\mu_i \nu_i} g_{\mu_i \nu_i}$ reduces to unit tensors ($g^{\mu_i \nu_i}$ is the inverse of $g_{\mu_i \nu_i}$) and we obtain:

$$F = \frac{1}{\sqrt{g'}} \sum_{k=1}^d dv_k \left(\frac{1}{k!} \nabla^{(\nu_1} \dots \nabla^{\nu_k)} F \right) \Big|_{\tau=\tau_0} \quad (4.17)$$

The last term in this summation as we showed in corollary 4.1, will be equal to joint probability i.e. $f^{\mu_1 \mu_2 \dots}$ tensor:

$$f^{\mu_1 \mu_2 \dots} = \frac{1}{\sqrt{g'} k!} \nabla^{(\mu_1} \dots \nabla^{\mu_k)} F \quad (4.18)$$

Corollary 4.4 In a flat manifold with Cartesian coordinates this equation reduces to (4.1)

Proof: regarding the definition of contravariant derivative $\nabla^\nu = g^{\mu\nu} \nabla_\mu$ and replacing the covariant derivative in terms of connection $\Gamma_{\beta\gamma}^\alpha$ (Christoffel symbol) we obtain:

$$\nabla^\nu = g^{\mu\nu} (\partial_\mu + [\text{terms in } \Gamma]) \quad (4.19)$$

In a flat space with Cartesian coordinate, all terms of $\Gamma_{\beta\gamma}^\alpha$ vanishes and (4.19) reduces to:

$$\nabla^\nu = g^{\mu\nu} \partial_\mu = \partial^\nu \quad (4.20)$$

Substitution of $\nabla^{(\mu_1} \dots \nabla^{\mu_k)}$ with (4.20) and noting the commutation property of ordinary partial derivative and $g' = g^{-1} = 1$ in Cartesian coordinate gives:

$$\nabla^{(\mu_1} \dots \nabla^{\mu_k)} = \partial^{(\mu_1} \dots \partial^{\mu_k)} = k! \partial^{\mu_1} \dots \partial^{\mu_k} \quad (4.21)$$

Hence equation (4.18) reduces to (4.1)

5 Results

Corollary 5.1 Almost all physical tensors of rank 2 or more, can be interpreted as a joint probability density, or simply a variable density. Of course, this proposition is the inverse of the lemma 3.6. Eventually the tensors $f^{ijk\dots}$ as *probability density*, reminds us the physical tensors such as stress-energy tensor $T^{\mu\nu}$ with a definition of energy density or energy flow density in Riemannian curved space-time manifold. Therefore this type of tensors could exemplify a physical realization of paired joint probability densities on Riemannian space.

5.1 Generalized Lagrange metric

As an example for existence of a scalar F in physics whose second partial derivative leads to metric tensor g_{ab} , we recall the so-called generalized Lagrange metric [7]:

$$g = \frac{\partial^2 \mathcal{L}}{\partial u^a \partial u^b} dx^a \otimes dx^b \quad (5.1)$$

Here the metric tensor stated as:

$$g_{ab} = \frac{\partial^2 \mathcal{L}}{\partial u^a \partial u^b} \quad (5.2)$$

Therefor the Lagrangian \mathcal{L} resembles the F in equation (4.1). Moreover for a locally flat manifold there always exist a metric by definition of the form:

$$g_{ij} = \frac{\partial^2 U}{\partial u^i \partial u^j} \quad (5.3)$$

These manifolds well-known as Hessian manifold and in the general case could be characterized by:

$$g_{ij} = \nabla_i \nabla_j U \quad (5.4)$$

Where the partial derivatives in the previous equation have been replaced by covariant derivatives as a generally accepted definition for Hessian. U stands for a scalar usually well-known as potential.

6 Discussion

In this short article a general form of joint probability density compatible with Riemannian parametric manifold has been introduced. We proved the contravariant tensor properties of this general form. Approach to these results facilitated by introducing a binary data matrix that collects the parametric information of a system of particles on a Euclidean *particle oriented coordinate* where the joint probability densities identified as a new definition *generalized inner products*. Extending the result to Riemannian sub-manifolds results in the general form of probability densities on Riemannian manifolds. It has been shown the reduction of this generalized joint probability density to the usual form of iterative partial derivative of cumulative probability function. Compatibility of the results with physical tensors such as stress- energy tensor has been verified.

References

- 1) Sirinivasan, S.K.; Mehata.K,M: Probability and Random Process, 2nd ed.; Tata McGraw-Hill ,1981; pp. 146.
- 2) Robinson, E.A.: Probability theory and application, ,Springer Science, 1985; pp. 144
- 3) Amari, Shun-ichi, and Hiroshi Nagaoka. :Methods of Information Geometry, Translations of Mathematical Monographs (2007).
- 4) Hendriks, Harrie. "Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions. *The Annals of Statistics* 1990: 832-849.
- 5) Burbea, J., and C. Radhakrishna Rao. "Differential metrics in probability spaces." *Probab. Math. Stat* 3 1984: 241-258.
- 6) Rao, C. Radhakrishna (1954); On the use and interpretation of distance functions in statistics. Bull. Inst. Inter. Statist. 34, 90-100.
- 7) Punzi, R. Wohlfarth, M.N.R.: Geometry and stability of dynamical systems, pp.12 arXiv: 0810.5060v1 [math-ph] 28 Oct (2008)

- 8) Viaclovsky, Jeff A. "Math 865, Topics in Riemannian Geometry." (2007)
- 9) Amari, Shun-ichi, and John Armstrong. "Curvature of Hessian manifolds." *Differential Geometry and its Applications* 33 , 2014: 1-12.
- 10) DasGupta, Anirban. *Fundamentals of probability: a first course*. Springer Science & Business Media, 2010.
- 11) Tsirulev, A. N. "Analytic continuation of tensor fields along geodesics by covariant Taylor series." *Theoretical and Mathematical Physics* 102.3, 1995: 245-250
- 12) Sathiapalan, B. "Gauge invariance and equations of motion for closed string modes." *Nuclear Physics B* 889 (2014): 261-298.



© 2016 by the author; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).