International Journal of Computational and Applied Mathematics & Computer Science
DOI: 10.37394/232028.2024.4.6

Joint Probability Densities on Riemannian Manifolds are Symmetric Tensor Densities

MANOUCHEHR AMIRI
Tandis Hospital, Tehran, IRAN

Abstract: This paper presents the tensor properties of joint probability densities on a Riemannian manifold. Initially, we develop a binary data matrix to record the values of a large number of particles confining in a closed system at a certain time in order to retrieve the joint probability densities of related variables. By introducing the particle-oriented coordinate and the generalized inner product as a multi-linear operation on the basis of this coordinate, we extract the set of joint probabilities and prove them to meet covariant tensor properties on a general Riemannian space of variables. Based on the Taylor expansion of scalar fields in Riemannian manifolds, it has been shown that the symmetrized iterative covariant derivatives of the cumulative probability function defined on Riemannian manifolds also give the set of related joint probability densities equivalent to the aforementioned multi-linear method. We show these covariant tensors reduce to classical ordinary partial derivatives in ordinary Euclidean space with Cartesian coordinates and give the formal definition of joint probabilities by partial derivatives of the cumulative distribution function. The equivalence between the symmetrized covariant derivative and the generalized inner product has been concluded. Some examples of well-known physical tensors clarify that many deterministic physical variables are presented as tensor densities with an interpretation similar to probability densities.

Keywords: Joint probability density, Tensor Density, Riemannian manifold, Symmetrized Covariant derivative, Taylor expansion

Received: July 17, 2023. Revised: May 19, 2024. Accepted: June 15, 2024. Published: July 16, 2024.

1 Introduction

The current classical theory of probability distributions is based on the presumptive Euclidean space of parameters and random variables. In this sense, the joint probability densities are derived from sequential partial derivatives of the corresponding cumulative distribution function [7],[12],[15],[17]. 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 [1],[2],[9],[13]. Moreover using differential geometry, some variations of information theory having been devised on non-Euclidean (Riemannian) metric spaces by Fisher, Rao, Amari and others [2], [4]. 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 that first introduced by Fisher and Rao, is the so-called Fisher information metric [14]. Applying the Riemannian metric to define the basic concepts in statistics such as mean and covariance matrix of random variable, has also been introduced in other works [3], [9] and [13]. The basic need for a precise formalism for joint probability densities on non-Euclidean spaces appears in the problems related to diffusion phenomena on surfaces such as spheres and other non-Euclidean surfaces [10],[11],[5]. In this short article, we introduce a new model based on binary data arrays and matrices to analyze the variables distributions of a system of particles on the Riemannian manifold. A cloud of a large number of particles at a certain time is considered in the space of variables without analysis of its dynamical evolution and quantum physics uncertainties restrictions. The coordinates of variables in this space, is divided to infinitesimal intervals while each interval labelled by an ordered integer number. The variables of each particle occupies just one infinitesimal interval on each variable coordinate x^ν labeled by some integer number i. For any specified infinitesimal interval on each coordinate, there is a specific array with binary entries $\{0, 1\}$ that determines the particles whose variable x^ν is restricted to this interval. Based on this background, in Section (2) we develop the concept of particle oriented coordinates which span a flat Euclidean space on which we embed Riemannian manifolds of variable coordinates. The collection of all binary arrays of all variables yields a binary matrix containing the entire information of the particles system. By introducing a generalized inner product for a set of vectors, joint probability densities of variables are calculated as inner product of the binary arrays that stands for some vectors in cotangent space at each point on manifold and consequently the tensor density properties of joint probability density are proved. In Section (3) based on the tensor density property of the joint probabilities of variables at any point in the manifold, we present a new definition of joint probability densities by symmetrized covariant derivative of the cumulative distribution function. A new method to connect concepts in continuous and discrete probability theory and a novel interpretation of covariant derivative by generalized inner product has been proposed. In Section (4) some examples are presented that reveal the tensor density properties of famous physical entities.

Definition 1. The space (manifold) $\mathcal{M}$ of variables spanned by coordinates x^ν with $1 \leq \nu \leq d$ where d is the dimension of manifold. So the number of involved independent variables (degree of freedom) is d .The variable space $\mathcal{M}$ in present article, generally presumed to be a Riemannian manifold with local coordinates x^ν and basis vectors $e_\nu = \frac{\partial}{\partial x^\nu}$ at any point $p \in \mathcal{M}$. If we divide the coordinate x^ν into m_ν small intervals $\Delta x^\nu(i)$, while m_ν is a large number, then integer i refers to the ith interval of this parameter. This means that i stands for the coordinate value of point p along x^ν and ranges between 0 and a large integer m_ν:

$$1 \leq i \leq m_\nu \tag{1}$$

The whole manifold includes all coordinates x^ν and associated basis vectors. The overall number of intervals reads as follows:

$$m = \sum_{\nu=1}^{d} m_\nu \tag{2}$$

In this setting the vector space $\mathcal{M}$ is a lattice space where the coordinates of points on manifold specified by d digit numbers, therefor the related field of $\mathcal{M}$ will be $\mathbb{Z}$. The limit manifold $m_\nu \to \infty$ $\mathcal{M}$, is smooth.

Definition 2. Regarding the definition of cumulative distribution or joint distribution function (CDF) [7], [15]. We may define a function $F(x^1, x^2, ..., x^d)$ at any point $P(x^1, x^2, ..., x^d) \in \mathcal{M}$ as follows:

$$\begin{gathered} 0 \leq F \leq 1 \\ \lim_{x^i \to \infty} F = 1, \quad 1 \leq i \leq d \\ \lim_{x^i \to -\infty} F = 0, \quad 1 \leq i \leq d \end{gathered} \tag{3}$$

Axiom For a system consisting of a large number of particles, there is a smooth and differentiable probability density function in the manifold $\mathcal{M}$ that yields the density of particles in any volume element dV in the manifold. This postulate is consistent with the accepted postulates of the kinetic theory of gases and the Maxwell-Boltzmann distribution.

2 Introducing Particle Dependent Coordinates

Assume a system consisting of a large number N of particles confined in an interval of space-time. Taking into account of such a system of particles, brings us the advantage of choosing a sufficient huge number of particles, moreover we could substitute particles by any kind of systems defined by their arbitrary points in parametric space. Suppose that a set of independent parameters labeled by ν is to be considered in a small interval of time $\Delta\tau$. 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 are defined as in Definition 1, denoted by $\Delta x^\nu(i)$ where i stands for the ordered location number of an intervals on the coordinate x^ν:

$$x^\nu(i-1) < x^\nu(i) < x^\nu(i+1) \tag{4}$$

So the variable's value of each particle falls in just one of these intervals.

Definition 3. Let the basis vectors $\varepsilon_1 = (1, 0, 0, ..)$, $\varepsilon_2 = (0, 1, 0, ..)$, .. $\varepsilon_N = (0, 0, 0, .., 1)$ span a vector space $\mathcal{V}$ of the dimension N over the field $\mathbb{Z}$, where N is the number of particles. $\mathcal{V}$ is regarded as a lattice Euclidean space endowed with Cartesian coordinates. The particles are labeled by ordered integer numbers; therefore, we set ε_1 as the basis for the first particle and ε_2 as the basis for the second particle, and so on. Here any particle specifies an independent (basis) coordinate with two possible values 0 and 1. Obviously these basis are orthogonal. We call these set of basis as *particle oriented coordinate* that as a coordinate chart is homeomorphic to a sub-space of Euclidean lattice space $\mathbb{Z}^N$. In this case the dual basis presentation coincides the same original basis i.e. $\varepsilon^{*i} = \varepsilon_i$. The dual basis could also be represented as a binary array like ; $\varepsilon^{*3} = (0, 0, 1, .., 0)$. At any point on the manifold $\mathcal{M}$ the basis $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, .., \frac{\partial}{\partial x^d}\right)$ spans the tangent space $T_P\mathcal{M}$. The basis $\varepsilon^{*\mu}$ acts on the basis $\partial_\nu = \frac{\partial}{\partial x^\nu}$ of tangent space at the point $P \in \mathcal{M}$ by the relation $\varepsilon^{*\mu}(\partial_\nu) = \delta^\mu_\nu$ because $\varepsilon^{*\mu}$ is specified to the variable x^ν and acts on ∂_ν to return the component "1" for the related particle, while its action on ∂_μ which is independent from ∂_ν returns "0". Consequently, the vectors $\varepsilon^{*\mu}$ are living in cotangent space $T^*_P\mathcal{M}$.

For each interval $\Delta x^{\nu}(i)$, there are some particles that their variable x^{ν} places in this interval. Let record the results of these outcomes in an array whose entries are 0 or 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 example, for the first particle, if its parameter value x^{ν} falls in the range of $\Delta x^{\nu}(i)$, it returns 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 a system of particles, e.g.

$$e_{\nu}(i) = (1, 0, 0, 1, 1, 0, 1, ...) \tag{5}$$

Each of these 0 and 1 connected to a specific particle in the system and the total number of entries equals the total number of particles N. In this example, the first entry 1 means that the value of the variable x^{ν} for the particle with label 1 is in the ith interval. Hence, these vectors are members of a vector space of dimension N. As defined in Definition 3, one may attribute to any particle an independent basis ε_{ν} for i th particle as

$$\varepsilon_{\nu} = (0, 0, 0, ..1, ..0, 0) \tag{6}$$

where only the i th entry takes the value 1. The result of projection of a set of ε_{ν} on the coordinate x^{ν} turns out the array $e_{\nu}(i)$ in equation (5). This means that the sum of basis ε_{ν} of all particles with the common value of x^{ν} yields the vector $e_{\nu}(i)$. For example, the vector in Equation (5), is the sum $e_{\nu}(i) = \varepsilon_1 + \varepsilon_4 + \varepsilon_5 + ...$

The vectors $e_{\nu}(i)$ at each point on x^{ν} belongs to a sub-space $\mathcal{N}$ of $\mathcal{V}$ defined in definition 3. Obviously, the dimension of $\mathcal{N}$ is d. At any point P on $\mathcal{M}$, if the tangent space spanned by $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, .., \frac{\partial}{\partial x^d}\right)$ denoted by $T_P\mathcal{M}$ and the tangent space at the same point on $\mathcal{N}$ denoted by $T_P\mathcal{N}$, then there is a one-to-one mapping between these spaces.

Remark. With respect to the Definition 3, the vectors $e^{*\nu}(i)$ as the sum of ε^{*i} at any point $P\left(x^1, x^2, ..., x^d\right) \in \mathcal{M}$ are functional at P with local coordinates $\left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, .., \frac{\partial}{\partial x^d}\right)$, thus belong to the dual space of $\mathcal{M}$ i.e. $e^{*\nu}(i) \in T^*_P\mathcal{M}$

Lemma 1. 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 4, columns of $\mathfrak{D}_{m_{\nu}\times N}$ carry just one entry 1 while other entries take 0. If the n th component of $e_{\nu}(i)$ is denoted by $[e_{\nu}(i)]_n$, then the inner product $\langle e_{\nu}(i), e_{\nu}(j)\rangle$ can be read as a sum over all particles:

$$\langle e_{\nu}(i), e_{\nu}(j)\rangle = \sum_{n=1}^{N} [e_{\nu}(i)]_n [e_{\nu}(j)]_n \tag{7}$$

The components $[e_{\nu}(i)]_n$ and $[e_{\nu}(j)]_n$ do not take the 1 value simultaneously, because a particle can not take two values on x^{ν}, thus it is easy to conclude that inner product $\langle e_{\nu}(i), e_{\nu}(j)\rangle$ vanishes for $i \neq j$

$$\langle e_{\nu}(i), e_{\nu}(j)\rangle = \delta_{ij} \tag{8}$$

The orthogonality of these bases is proved. □

Definition 4. For a fixed ν, we define the matrix $\mathfrak{D}_{m_{\nu}\times N}$ with $e_{\nu}(i)$ as the ith row. m_{ν} as described in Definition 1, is the number of infinitesimal intervals in x^{ν}. Obviously each column of this matrix contains just one entry 1 and other entries are 0, because each column belongs to a particle which occupies just one value (interval) at $\Delta x^{\nu}(i)$.

Because of orthogonality, the set of $\{e_{\nu}(i) | 1 \leq i \leq m_{\nu}\}$ for a fixed ν span a tangent space $T^{\nu}_P\mathcal{N}$ which is in one to one mapping with tangent $T^{\nu}_P\mathcal{M}$ with coordinates $x^{\nu}(i)$.

$$Span\{e_{\nu}\} = T^{\nu}_P\mathcal{N} \tag{9}$$

$$T^{\nu}_P\mathcal{N} \subset T_P\mathcal{N} \tag{10}$$

The manifold $\mathcal{N}$ is a lattice Euclidean space, but the manifold $\mathcal{M}$ as a lattice space is not necessarily flat and may be endowed by a general Riemannian metric.

The probability density of particles within interval $dx^{\nu}(i)$ which will be denoted by $f_{\nu}(i)$, is proportional to the ratio of total number of entry "1" in the $e_{\nu}(i)$, denoted by $n_{\nu}(i)$, to total particles number N :

$$f_{\nu}(i)\, dx^{\nu}(i) = \frac{n_{\nu}(i)}{N} \tag{11}$$

It is straightforward to conclude that $n_{\nu}(i)$ equals the inner product $\langle e_{\nu}(i), e_{\nu}(i)\rangle$. In an infinitesimal limit of $\Delta x^{\nu}(i)$ the basis $e^{*\nu}(i)$ as the dual vector of $e_{\nu}(i)$ approaches $dx^{\nu}(i)$ which belongs to the cotangent (dual) space basis of $T^{*\nu}_P\mathcal{M}$ while it carries the information about that interval in the system considered. With respect to coordinate transformations $x^{\nu} \to \bar{x}^{\nu}$ in variable space, we apply the rule of $dx^{\nu}(i)$ transformation with summation on ν:

$$d\bar{x}^{\mu} = \frac{\partial \bar{x}^{\mu}}{\partial x^{\nu}} dx^{\nu} \tag{12}$$

In this equation and all tensor equations in the next sections, we use Einstein's summation convention on repeated covariant and contravariant indices.

2.1 Binary Data Matrix

Combining all row vectors $e_\nu(i)$ for all d variables of N particles gives a matrix $\mathfrak{D}_{m\times N}$ with m rows defined in equation (2) and N as the total number of particles. This matrix itself is also the combination of d matrices each for a specific variable. As an example if for a variable 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 $\mathfrak{D}_{m\times N}$ where $e^{*\nu}(i)$ constitute the set of dual basis for coordinates x^ν at any location i. The columns of $\mathfrak{D}_{m_\nu\times N}$ carries just one entry 1 , because each column corresponds to one particle that its variable's value falls in just one of intervals Δx^ν , namely $\Delta x^\nu(i)$. Obviously combining all $\mathfrak{D}_{m_\nu\times N}$ results in the data matrix $\mathfrak{D}_{m\times N}$. As we showed in Equation (9) the set of $\{e_\nu(i)\}$ spans a tangent space $T_P\mathcal{N}$ at a point $P\in\mathcal{N}$.

Remark. The joint probability density of two variables $x^\nu(i)$ and $x^\mu(j)$ is proportional to the number of particles that simultaneously have the same parameter values of $x^\nu(i)$ and $x^\mu(j)$ or equivalently are confined in $\Delta x^\nu(i)$ and $\Delta x^\nu(j)$ where i and j indicate the values of the coordinates x^ν and x^μ (i.e. coordinate values). If this number presented by $n^{\mu\nu}$ the exact form of joint probability density reads as:

$$f_{\mu\nu}dV=\frac{n_{\mu\nu}}{N}\quad \mu\neq\nu \tag{13}$$

$dV=dx^\mu(i)\,dx^\nu(j)$ is the volume element.

It is noteworthy to remind that for $\mu=\nu$ the joint density $f_{\mu\nu}$ and $n_{\mu\nu}$ reduce to f_μ and n_μ respectively as will be shown in Lemma 3 .

Definition 5. We define the generalized inner (scalar) product $\odot$ for vectors $U,V,W,..$ in a vector space $\mathbb{V}$ with orthogonal local coordinates as a multi-linear map:

$$\Phi:\mathbb{V}^m\to\mathbb{R}\quad;\quad U\odot V\odot W...=\sum_{n=1}^{m}u_nv_nw_n... \tag{14}$$

Where u_n, v_n, .. are $n-th$ components of U, V, .. respectively. This is simply a generalization of inner products in the usual definition. For inner product of two basis vector in Cartesian coordinate on Euclidean tangent or cotangent space, results in metric tensor. As an example, the inner product for e_i and $e_j\in T_P\mathcal{M}$ results in the metric tensor $g_{ij}=\left\langle\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right\rangle$

Lemma 2. Generalized scalar operation $\odot$ is linear and commutative.

$$\begin{aligned}U\odot V&=V\odot U\\ \lambda_1U\odot\lambda_2V&=\lambda_1\lambda_2V\odot U\\ (U+V)\odot W&=U\odot W+V\odot W\end{aligned} \tag{15}$$

Proof. Due to Definition 5, it is straightforward to derive the equations of linear and commutative properties. □

If the vector U is a binary vector, the idempotent property also holds:

Lemma 3. For inner product $\odot$ of k vectors when $k\geq 3$, the repeated vectors reduces to one vector

$$U\odot V\odot V\odot W=U\odot V\odot W \tag{16}$$

Proof. Applying definition in equation (14) and commutative property of $\odot$ proves the equation (16). □

2.2 Joint probability densities as tensor densities

Theorem 4. *The joint probability density $f_{\mu\nu\xi..}(i,j,k,..)$ for particles with common coordinate values $x^\mu(i)$, $x^\nu(j)$, $x^\xi(k)$... can be given by the generalized inner product of the basis $\{e_\mu(i),e_\nu(j),e_\xi(k),..\}$ via the equation:*

$$\begin{aligned}&f_{\mu\nu\xi..}(i,j,k,..)\,dV(i,j,k...)=\\&\lim_{N\to\infty}\frac{1}{N}e_\mu(i)\odot e_\nu(j)\odot e_\xi(k)\,..\end{aligned} \tag{17}$$

Proof. By using the equation (14) for basis $\{e_\mu(i),e_\nu(j),e_\xi(k),..\}$ and by omitting the location indices i , j , k , .. we have:

$$e_\mu\odot e_\nu..=\sum_{n=1}^{m}(e_\mu)_n(e_\nu)_n\,.. \tag{18}$$

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 (18) reduces to the number of particles whose parameter values simultaneously are located in the intervals i , j , k , .. in the corresponding x^μ,x^ν,x^ξ,.. coordinates. This number is denoted as $n_{\mu\nu\xi}$. Normalization of $n_{\mu\nu\xi}$ as defined in equations (13) and (17) by $1/N$ yields the ratio of this number to total number of particles and consequently gives the joint probability $f_{\mu\nu\xi..}$ of particles with common coordinate values $x^\mu(i)$, $x^\nu(j)$ .. .

$$f_{\mu\nu\xi..}dV(i,j,k...)=\frac{n_{\mu\nu\xi..}}{N}=\frac{1}{N}e_\mu(i)\odot e_\nu(j)\odot e_\xi(k)\,.. \tag{19}$$

□

Due to the mentioned axiom, the joint probability $f_{\mu\nu\xi..}dV(i,j,k...)$ and $\frac{n_{\mu\nu\xi..}}{N}$ are smooth differential forms at the limit $N\to\infty$ and $dV\to 0$. The equation (19) reveals a specific configuration or 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 can be represented by such a specific set $\{e_\mu(i),e_\nu(j),e_\xi(k)\,..\}$ for all points on manifold $\mathcal{M}$ and vice versa. Evidently the order of indices $\mu,\nu\,\xi$ .., does not affect on the related joint probability density. The example of this symmetry could be found in symmetric properties of metric tensors $g_{ij} = \langle e_i, e_j \rangle$ or $g^{ij} = \langle e^{*i}, e^{*j} \rangle$ with respect to their lower and upper indices, respectively. We show the tensor properties of $f_{\mu\nu\xi..}$ in the next theorem.

Lemma 5. $\frac{n_{\mu\nu\xi..}}{N}$ are covariant tensors.

Proof. Let $V_\nu = T_P^\nu \mathcal{N} \subset T_P \mathcal{N}$ and $V^{*\nu} = T_P^{*\nu} \mathcal{N} \subset T_P^* \mathcal{N}$. From equation (19) we have the following.

$$\frac{n_{\mu\nu\xi..}}{N} = \frac{1}{N} e_\mu (i) \odot e_\nu (j) \odot e_\xi (k) .. \tag{20}$$

Since $e_\mu \in V_\mu$, the equation (19) is a map ϕ from $V_\mu \times V_\nu \times V_\xi ..$ to $\mathbb{R}$ or in a brief notation:

$$\phi : V^d \to \mathbb{R} \tag{21}$$

Respect to linear properties of $\odot$ in equations (14) ϕ is a multi-linear map with the following property:

$$\varphi (e_1, ..., ae_i + b\varepsilon_i, ..., e_d) = \\ e_1 \odot ... \odot ae_i \odot ... \odot e_d + e_1 \odot ... \odot b\varepsilon_i \odot ... \odot e_d \tag{22}$$

$$\begin{gathered} \varphi (e_1, ..., ae_i + b\varepsilon_i, ..., e_d) = \\ a\varphi (e_1, ..., e_i, ..., e_d) + b\varphi (e_1, ..., \varepsilon_i, ..., e_d) \end{gathered} \tag{23}$$

There is a one to one map between local tangent vectors $\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, .., \frac{\partial}{\partial x^d}$ on $\mathcal{M}$ and tangent vectors $e_1, e_2, .., e_d$ on $\mathcal{N}$ at any point P. Since $\mathbb{Z}$ is the common field of both vector spaces, therefore the tangent spaces $T_P\mathcal{N}$ and tangent space $T_P\mathcal{M}$ are isomorphic. The term $\frac{n_{\mu\nu\xi..}}{N}$, due to the axiom in section (1) and equation (19) is a smooth differential form and as a multi-linear map defined in (20) is transformed as covariant tensor at any point $p \in \mathcal{M}$, with the rank $\leq d$ similar to equation (12). Thus, under the coordinate transformation $x^\nu \to \bar{x}^\nu$ we have:

$$\frac{\bar{n}_{\mu\nu\xi..}}{N} = \left(\frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\xi} .. \right) \frac{n_{\alpha\beta\gamma..}}{N} \tag{24}$$

□

Theorem 6. *The joint probability density $f_{\mu\nu\xi..}$ defined in (19) is a covariant tensor density of weight -1.*

Proof. With the coordinate transformation $x^\nu \to \bar{x}^\nu$, the equations (19) and (24) give:

$$\bar{f}_{\mu\nu\xi..} d\bar{V} = \frac{\bar{n}_{\mu\nu\xi..}}{N} = \left(\frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\xi} .. \right) \frac{n_{\alpha\beta\gamma..}}{N} \tag{25}$$

Respect to equation (19) we have:

$$\bar{f}_{\mu\nu\xi..} d\bar{V} = \left(\frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\xi} .. \right) f_{\alpha\beta\gamma..} dV \tag{26}$$

Due to the definition of volume elements dV and $d\bar{V}$, we obtain:

$$\bar{f}_{\mu\nu\xi..} = |J|^{-1} \left(\frac{\partial x^\alpha}{\partial \bar{x}^\mu} \frac{\partial x^\beta}{\partial \bar{x}^\nu} \frac{\partial x^\gamma}{\partial \bar{x}^\xi} .. \right) f_{\alpha\beta\gamma..} \tag{27}$$

□

Where $|J|$ denoted as determinant of Jacobian. With respect to the definition of the tensor density, $\bar{f}_{\mu\nu\xi..}$ in equation (27) is a tensor density of weight -1.

Remark. Taking into account the commutative property of $\odot$ from equations (15), the tensor density $f_{\mu\nu\xi..}$ is symmetric with respect to the covariant indices $\mu, \nu, \xi.$ as expected for a joint probability. On the other hand repeated covariant indices reduce to non-repeated indices as we showed in Lemma 3:

$$f_{\alpha\beta\beta\gamma..} = f_{\alpha\beta\gamma..} \tag{28}$$

3 Joint Probability Densities as Symmetrized Covariant Derivatives of Cumulative Distribution Function

In the context of probability theory, the joint probability density of multiple (random) variables which are defined on a flat Euclidean space equipped with Cartesian coordinate, presented as a sequence of partial derivative of cumulative distribution function [7],[12],[15] and [17]:

$$f_{ijk..} = \frac{\partial^n F}{\partial x^i \partial x^j \partial x^k ..} = (\partial_i \partial_j \partial_k ..) F \tag{29}$$

Where F stands for the cumulative distribution function (CDF) and $\partial_i = \frac{\partial}{\partial x^i}$. In this sequence of partial derivatives, repetition of indices is not allowed. Obviously $f_{ijk..}$ is not a tensor and does not meet the transformation requirement of tensors. Therefore, in a flat Euclidean space with Cartesian coordinates $(x^1, x^2, ..)$, the joint probability density under coordinate transformation $x^\nu \to \bar{x}^\nu$ will transform as [12]:

$$\bar{f}\left(\bar{x}^1, \bar{x}^2, .., \bar{x}^d\right) = |J|^{-1} f\left(x^1, x^2, .., x^d\right) \tag{30}$$

3.1 Taylor Expansion

$f_{\mu\nu\xi..}$ as a covariant tensor is full symmetric respect to covariant indices and fulfils this property of joint probability densities. Joint probability densities for cumulative distribution function on Euclidean flat manifolds, are derived by regular partial derivatives respect to the contravariant coordinates x^i .

Lemma 7. If a differentiable smooth scalar function ϕ is defined on a d-dimensional Euclidean space with Cartesian coordinates, then the Taylor expansion of ϕ around a point $\left(x_0^1, x_0^2, .., x_0^d\right)$ when for all ν, $(x^\nu - x_0^\nu) \to 0$ will read as follows:

$$\phi\left(x^1, x^2, .., x^d\right) = \phi\left(x_0^1, x_0^2, .., x_0^d\right) + \frac{\partial \phi}{\partial x^1}\left(x^1 - x_0^1\right) + \frac{\partial \phi}{\partial x^2}\left(x^2 - x_0^2\right) + .. + \frac{\partial^d \phi}{\partial x^1 \partial x^2 .. \partial x^d}\left(x^1 - x_0^1\right) .. \left(x^d - x_0^d\right) \tag{31}$$

Proof. For a function $\phi\left(x^1, x^2, .., x^d\right)$ defined on a d-dimensional space, at the limit $(x^\nu - x_0^\nu) \to 0$ if the order of $\left(x^1 - x_0^1\right) .. (x^\nu - x_0^\nu)$ be higher than d, it will be negligible. Therefor the expansion at this limit will close at the d th order of the partial derivatives. □

Let $\phi = F$ be the cumulative distribution function. Because at the limit $x^\nu \to x_0^\nu$, the term $\left(x^1 - x_0^1\right) .. \left(x^d - x_0^d\right)$ in Cartesian coordinates is the volume element, its coefficient in the last term is the joint probability density for variables d at $x^\nu \to x_0^\nu$, as shown in equation (31).

The Taylor expansion could be generalized to the Taylor expansion on a Riemannian manifold by means of *symmetric (symmetrized) covariant derivatives*. Symmetrization of the multiple covariant derivative (symmetrized covariant derivative) could be accomplished by routine symmetrized form [16],[19]:

$$T_{\mu\nu..\kappa} = \frac{1}{k!}\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)}F \tag{32}$$

This notation with bracket around indices is the abbreviation for the sum of all permutations over $1 \leq k \leq d$ indices. In this sequence of covariant derivatives, the repetition of indices is not allowed. Obviously $T_{\mu\nu..\kappa}$ is a tensor. In the Taylor expansion for F on a Riemannian manifold, the last symmetrized consecutive covariant derivatives is the coefficients of $\left(x^1 - x_0^1\right) .. \left(x^d - x_0^d\right)$ at the limit $x^\nu \to x_0^\nu$ where the coordinates x^i are the local coordinates of Riemannian manifold $\mathcal{M}$. At this limit the last term reads as:

$$T_{\mu\nu..\kappa}dx^1dx^2..dx^d = \frac{1}{d!}\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)}F dx^1dx^2..dx^d \tag{33}$$

By multiplying and dividing to $\sqrt{g}$ and taking into account the property $\nabla_\mu g_{\nu\xi} = 0$, we get:

$$T_{\mu\nu..\kappa}dx^1dx^2..dx^d = \frac{1}{\sqrt{g}d!}\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)}F\sqrt{g}dx^1dx^2..dx^d \tag{34}$$

In this equation the last term $F\sqrt{g}dx^1dx^2..dx^d$ is a scalar density which is invariant under integration on the possible domain and the term $\sqrt{g}dx^1dx^2..dx^d$ is invariant volume element. Since $\sqrt{g}$ is a tensor density of weight 1, the term :

$$\frac{1}{\sqrt{g}d!}\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)}F \tag{35}$$

is a tensor density of weight -1. Therefor respect to equation (27), and the fact that this term is the coefficient of invariant volume element, it equals to the joint probability density:

$$f_{\mu\nu..\kappa} = \frac{1}{\sqrt{g}d!}\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)}F \tag{36}$$

$f_{\mu\nu..\kappa}$ is symmetric with respect to the indices as expected for a joint probability density. Actually these covariant derivatives of any order remain as tensors and preserve the tensor properties of $f_{\mu\nu..\kappa}$. The equation (36) after contraction by metric tensors $g^{\alpha\mu}$ could be presented in the contravariant form. We define $g^{\alpha\mu}\nabla_\mu = \nabla^\alpha$ as "contravariant derivative" [18]. Because of the identity $\nabla_\xi g^{\alpha\mu} = 0$ and the fact that covariant derivative of tensor products obey the Leibniz rule ([8] page110), the derivative ∇_ξ and $g^{\alpha\mu}$ are commutative, thus after multiplying both side of equation (36) by $g^{\alpha\mu}g^{\beta\nu}..g^{\delta\kappa}$ we obtain:

$$f^{\alpha\beta..\delta} = \frac{1}{\sqrt{g}d!}\nabla^{(\,\alpha}\nabla^\beta..\nabla^{\delta\,)}F \tag{37}$$

This is the joint probability density defined in dual space of variables.

Regarding the definition of covariant derivative of the scalar F, for the first order covariant derivative we have:

$$\nabla_\nu F = \frac{\partial}{\partial x^\nu}F \tag{38}$$

Adding more symmetrized covariant derivatives leads to the terms that contains Christoffel symbols [20]:

$$\begin{aligned}\nabla_{(\,\mu}\nabla_{\nu\,)}F &= \frac{1}{2}\left(\nabla_\mu\nabla_\nu + \nabla_\nu\nabla_\mu\right)F \\ &= \nabla_\mu\nabla_\nu F = \partial_\mu\partial_\nu F - \Gamma^\xi_{\mu\nu}\partial_\xi F\end{aligned} \tag{39}$$

Where, the identity $\nabla_\mu\nabla_\nu F = \nabla_\nu\nabla_\mu F$ is used in equation (39).

Lemma 8. In a flat Euclidean manifold with Cartesian coordinates, equation (36) reduces to (29).

Proof. In an Euclidean space with Cartesian coordinates all terms of $\Gamma^\xi_{\mu\nu}$ vanish and ∇_μ reduces to ∂_μ:

$$\nabla_{(\,\mu}\nabla_\nu..\nabla_{\kappa\,)} = d!\partial_\mu\partial_\nu..\partial_\kappa \tag{40}$$

Taking into account the identity $g = 1$ for this case, the equation (36) becomes:

$$f_{\mu\nu..\kappa} = \partial_\mu \partial_\nu .. \partial_\kappa F \tag{41}$$

□

Therefore, the joint probability density in Equation (29) is a special case of Equation (36).

3.2 Equivalence of Symmetrized Covariant Derivative and Generalized Inner Product

In previous sections the joint probability densities have been derived by two distinct methods 1) Symmetrized covariant derivatives on Riemannian manifold 2) Generalized inner product defined on "Particle-Oriented coordinates". This conveys the idea of equivalence of these two operation on basis vectors of connected and discrete spaces respectively. Based on Equations (19) and (36) this equivalence can be shown by the following notation:

$$\begin{aligned} \nabla_\mu F &\sim e_\mu \\ \nabla_{(\,\mu}\nabla_{\nu\,)} F &\sim e_\mu \odot e_\nu \\ \nabla_{(\,\mu}\nabla_\nu \nabla_{\xi\,)} F &\sim e_\mu \odot e_\nu \odot e_\xi \end{aligned} \tag{42}$$

These relations suggest a new method to connect the concepts in discrete and continuous probability theory and a novel interpretation of covariant derivatives in differential geometry.

4 Examples

In this section, some examples of the application of joint probability densities in physics and engineering are presented. Actually the joint probability densities $f^{\mu\nu\xi..}$ and $f_{\mu\nu\xi..}$ as tensor densities, remind us the physical tensors such as stress-energy tensor $T^{\mu\nu}$ which stands for definitions of energy density, momentum density or energy flow density in Riemannian curved space-time manifold. Therefore these type of tensors could exemplify a physical realization of joint probability densities of two variables on the Riemannian space. the following example reveals the physical interpretation of joint probability densities as tensor density in various field of physics.

4.1 Example 1

Four-current density has many applications in engineering and its transformation under $x^\nu \to \bar{x}^\nu$ on a manifold with metric tensor $g_{\mu\nu}$ is represented as a four-vector density by [6]:

$$\bar{\mathrm{j}}^\mu \equiv |J|^{-1} \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \mathrm{j}^\alpha \tag{43}$$

where j^μ is a contravariant tensor density of rank -1.

4.2 Example 2

The stress-energy tensor T^{ij}, with $i, j \in \{0, 1, 2.3\}$ that has been introduced in the context of general relativity, as a contravariant tensor of rank 2 [21] , $\mu, \nu \in \{1, 2.3\}$, comprises of five physically different components ; T^{00} as energy density, $T^{\mu 0}$ as energy flux density, $T^{0\mu}$ as momentum density, $T^{\mu\mu}$ as pressure and $T^{\mu\nu}$ with $\mu \neq \nu$ as shear stress density. Thus, the stress-energy tensor $T^{\mu\nu}$ is a physical example of a joint probability density tensor in a deterministic case of physics such as general relativity. This fact conveys the idea that many deterministic physical variables have a fundamentally probabilistic origin.

5 Conclusion

This article introduces a general form of the joint probability density of variables that are defined on a Riemannian manifold. It is shown that the joint probability densities on Riemannian manifold transforms as tensor densities of weight -1. Approach to these results facilitated by introducing a binary data matrix that collects the variables information of a system of particles on a lattice Euclidean space embedded by the particle oriented coordinates where the joint probability densities identified as a new definition of generalized (multi-linear) inner products of basis vectors. By this method the tensor density properties of joint probability is proved. Based on Taylor expansion of scalar field in a Riemannian manifolds,it has been shown that the symmetrized iterative covariant derivatives of cumulative probability function defined on the Riemannian manifold also give the set of related joint probability densities equivalent to the generalized inner products method. As an outcome, the equivalence of symmetrized iterative covariant derivatives and multilinear inner product is proved. It has been shown that, in Euclidean space of variables with Cartesian coordinates, the generalized joint probability density reduces to the usual form of iterative partial derivative of cumulative function. A new method to connect concepts in continuous and discrete probability theory and a novel interpretation of covariant derivative by generalized inner product has been proposed. Some examples of well-known physical tensors convey us that many deterministic physical variables may have fundamentally a probabilistic origin that through the future work on this subject will be more clarified.

Declarations

- The author confirms sole responsibility for the study conception and manuscript preparation.
- No funding was received for conducting this study.

- The author declares that he has no conflict of interest.
- Due to the nature of the research, supporting data is not available.

References

[1] S.-i. Amari and J. Armstrong. Curvature of hessian manifolds. *Differential Geometry and its Applications*, 33:1–12, 2014.

[2] S.-i. Amari and H. Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2000.

[3] A. Bhattacharya and R. Bhattacharya. Statistics on riemannian manifolds: asymptotic distribution and curvature. *Proceedings of the American Mathematical Society*, 136(8):2959–2967, 2008.

[4] J. Burbea and C. Radhakrishna Rao. Differential metrics in probability spaces. *Probability and mathematical statistics*, 3(2):241–258, 1984.

[5] K. Burrage, P. M. Burrage, and G. Lythe. Effective numerical methods for simulating diffusion on a spherical surface in three dimensions. *Numerical Algorithms*, 91(4):1577–1596, 2022.

[6] F. Cabral and F. S. Lobo. Electrodynamics and spacetime geometry: foundations. *Foundations of Physics*, 47:208–228, 2017.

[7] A. DasGupta. *Fundamentals of probability: A first course*. Springer Science & Business Media, 2010.

[8] P. Hajicek. *An introduction to the relativistic theory of gravitation*, volume 750. Springer Science & Business Media, 2008.

[9] H. Hendriks. Nonparametric estimation of a probability density on a riemannian manifold using fourier expansions. *The Annals of Statistics*, pages 832–849, 1990.

[10] C.-W. Huang, M. Aghajohari, J. Bose, P. Panangaden, and A. C. Courville. Riemannian diffusion models. *Advances in Neural Information Processing Systems*, 35:2750–2761, 2022.

[11] O. Orrick, M. Yang, C. Batchelor-McAuley, and R. G. Compton. Characterising fickian diffusion on the surface of a sphere. *Journal of Electroanalytical Chemistry*, 900:115738, 2021.

[12] A. Papoulis and S. Unnikrishna Pillai. *Probability, random variables and stochastic processes*. McGraw-Hill, 2002.

[13] X. Pennec. *Probabilities and statistics on riemannian manifolds: A geometric approach*. PhD thesis, INRIA, 2004.

[14] C. R. Rao. On the use and interpretation of distance functions in statistics. *Bulletin of the International Statistical Institute*, 34:90–100, 1954.

[15] E. A. Robinson et al. *Probability theory and applications*. Springer, 1985.

[16] B. Sathiapalan. Gauge invariance and equations of motion for closed string modes. *Nuclear Physics B*, 889:261–298, 2014.

[17] S. K. Srinivasan and K. Mehata. Probability and random processes. *(No Title)*, 1978.

[18] J. L. Synge and A. Schild. *Tensor calculus*, volume 5. Courier Corporation, 1978.

[19] A. N. Tsirulev. Analytic continuation of tensor fields along geodesics by covariant taylor series. *Theoretical and Mathematical Physics*, 102:245–250, 1995.

[20] J. A. Viaclovsky. Math 865, topics in riemannian geometry, 2007.

[21] R. M. Wald. *General relativity*. University of Chicago press, 2010.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The author contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The author has no conflict of interest to declare that is relevant to the content of this article.