EXTRACTA MATHEMATICAE Vol. 9, Núm. 2, 119 – 123 (1994)

Differentiation of Measures *

RICARDO FARO, JUAN A. NAVARRO AND JUAN SANCHO

Departamento de Matemáticas, Universidad de Extremadura, 06071-*Badajoz, Spain*

AMS *Subject Class.* (1991): 28C99. *Received July* 8, 1994

0. INTRODUCTION

The notation $\mathrm{d}\mu$, referred to a measure μ, often appears in many texts on measure theory. This notation is used by its very suggesting character, but without any previously defined notion of differential of a measure. For example, the Radon–Nikodym derivative is frequently denoted by a quotient $\mathrm{d}\mu/\mathrm{d}\lambda$, but no meaning is given to the individual terms $\mathrm{d}\mu$ and $\mathrm{d}\lambda$.

Our aim is to overcome this deficiency, defining the differential of a measure at a point in the realm of $\mathscr{C}^1$-differentiable manifolds. This new concept will lead us to a better understanding of the Radon–Nikodym derivative and it could be useful in the theory of conditional probability.

1. DEFINITION OF THE DIFFERENTIAL OF A MEASURE

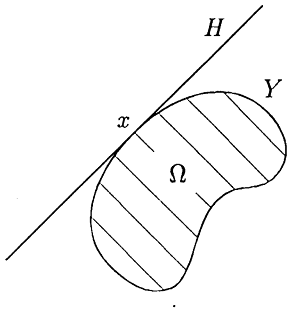
Let μ be a measure defined on the Borel subsets of a smooth manifold X. Our aim is to define the *differential* $\mathrm{d}_x\mu$ of μ at any point x of X, so that $\mathrm{d}_x\mu$ should reflect the infinitesimal behaviour of μ around x. Hence $\mathrm{d}_x\mu$ will be a measure of infinitesimal figures at x; that is to say, of figures in the tangent vector space T_xX of X at x. Briefly, $\mathrm{d}_x\mu$ should be a measure on T_xX.

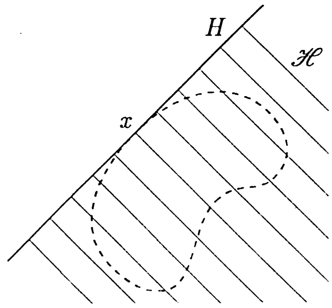
For the sake of simplicity, let us assume that our differentiable manifold X is a n-dimensional real vector space E and the given point x is the origin of E. In this case the tangent vector space T_xX is canonically identified with E, so that we may consider that μ and $\mathrm{d}_x\mu$ are defined on the same space E.

It may be convenient to see an example. Let Ω be a region of E limited by an hypersurface Y and let x be a point of Y. Given a Lebesgue measure m on E, we consider the measure μ_Ω defined by setting

$$\mu_\Omega(A) = m(A \cap \Omega) \ .$$

*A full version will appear in *Archiv der Math.*





From an intuitive viewpoint, it is clear that the infinitesimal behaviour of μ_Ω around x is determined by the tangent hyperplane H to the hypersurface Y at x. That is to say, whatever the definition of the differential would be, it is intuitively clear that the differential of μ_Ω at x should be the measure $\mu_{\mathscr{H}}$, where $\mathscr{H}$ is the half-space limited by the hyperplane H.

These examples suggest that a natural definition of the differential of a measure μ on E at the point $x = 0$ could be provided by the following formula:

$$(*) \qquad (\mathrm{d}_x\mu)(A) = \lim_{t\to 0^+} \mu(tA)/t^n$$

But it is not difficult to give examples where the differential is "evident" while the formula (*) is not valid for any Borel set A. It is too strong to impose the equality (*) for any Borel set A.

Let us see the above formula (*) from another viewpoint. Remark that $\mu_t(A) = \mu(tA)/t^n$ defines a measure on E for every $t \in \mathbb{R}_+$. The problematic formula (*) suggests that the differential $\mathrm{d}_x\mu$ should be, in some sense, the limit of the measures μ_t as $t \longrightarrow 0^+$. The sense of this limit is provided by the theorem of Riesz: Measures that are finite on compact sets form the dual space of the space of continuous functions with compact support. Hence, it is natural to consider the weak topology on the measures (finite on compact sets), so that we shall say that a sequence of measures $\{\mu_i\}$ converges to a measure μ if

$$\int f\,\mathrm{d}\mu_i \quad\longrightarrow\quad \int f\,\mathrm{d}\mu$$

for any continuous function f with compact support. These remarks lead us to the following definition:

DEFINITION. Let μ be a measure on E finite on compact sets and let x be any point of E. We say that μ is *differentiable* at x if there exists a measure $\mathrm{d}_x\mu$ such that $\mathrm{d}_x\mu = \lim_{t\to 0^+} \mu_t$, where μ_t are the measures defined by

$$\mu_t(A) = \mu(x + tA)/t^n \ .$$

This measure $\mathrm{d}_x\mu$ is said to be the *differential* of μ at x.

This definition may be easely modified so as to give it a sense for measures defined on an open subset of E. From now on, *any measure will be assumed to be finite on compact sets.*

The above definition is related to the intuitive equality (*) by the following result:

PROPOSITION. *Let μ be a measure on E. If μ is differentiable at a point x, then we have*

$$(d_x\mu)(A) = \lim_{t\to 0^+} \mu(x + tA)/t^n$$

for any bounded Borel set A such that $(d_x\mu)(\partial A) = 0$.

2. Properties of the Differential

The differential of measures has the following elementary properties:

1. The differentiability of measures is a local concept: If two measures μ, λ coincide on a neighbourhood of x, then μ is differentiable at x if and only if so is λ, and in this case $\mathrm{d}_x\mu = \mathrm{d}_x\lambda$.

2. The differential $\mathrm{d}_x\mu$ is an homogeneous measure: $(\mathrm{d}_x\mu)(tA) = t^n(\mathrm{d}_x\mu)(A)$, $t \in \mathbb{R}_+$.

3. Let μ, λ be differentiable measures at x. Then $a\mu + b\lambda$ is differentiable at x for any $a, b \in \mathbb{R}_+$ and

$$\mathrm{d}_x(a\mu + b\lambda) = a\,\mathrm{d}_x\mu + b\,\mathrm{d}_x\lambda \ .$$

4. Any Lebesgue measure m on E is differentiable at any point x and $\mathrm{d}_x m = m$.

5. Let $f \geq 0$ be an integrable function, continuous at x. If a measure μ is differentiable at x, then so is $f\mu$ and

$$\mathrm{d}_x(f\mu) = f(x)\,\mathrm{d}_x\mu \ .$$

3. Formula of Change of Variables

The following result is a differential statement of the formula of change of variables in the integration theory.

THEOREM. *Let $\varphi : U \longrightarrow V$ be a $\mathscr{C}^1$-diffeomorphism between open subsets of $\mathbb{R}^n$ and let μ be a measure on U. If μ is differentiable at a point x of U, then $\varphi(\mu)$ is differentiable at $y = \varphi(x)$ and we have*

$$\varphi'(d_x\mu) = d_y(\varphi(\mu)) \ ,$$

where φ' stands for the derivative of φ at x, and $\varphi(\mu)$ denotes the following measure: $\varphi(\mu)(A) = \mu(\varphi^{-1}(A))$.

This theorem enables us to define the differential $\mathrm{d}_x\mu$ of a measure μ (finite on compact sets) on a $\mathscr{C}^1$-differentiable manifold. In fact, if we choose a coordinate system $(u_1, \ldots, u_n)$ defined on an open neighbourhood U of x, it defines a diffeomorphism φ of U onto an open subset $\bar{U}$ of $\mathbb{R}^n$. Now, μ is said to be differentiable at x whenever $\bar{\mu} = \varphi(\mu)$ is differentiable at $\bar{x} = \varphi(x)$ and, in such case, the differential of μ at x is a measure on T_xX defined by the equality $\mathrm{d}_x\mu = (\varphi')^{-1}(\mathrm{d}_{\bar{x}}\bar{\mu})$, where $\varphi' : \mathrm{T}_xX \longrightarrow \mathrm{T}_{\bar{x}}\mathbb{R}^n$ denotes the derivative of φ at x. The above theorem just states that these definitions do not depend on the choosen coordinate system.

4. Points of Differentiability

Let X be a $\mathscr{C}^1$-differentiable manifold and let m be the measure defined by some riemannian metric on X. A Borel subset A is said to be *null* if $m(A) = 0$. It is easy to prove that this definition does not depend on the choosen riemannian metric.

THEOREM. *Any measure on X is differentiable at every point, except for a null set. Moreover, $\mathrm{d}_x\mu$ and d_xm are proportional almost everywhere.*

As a consequence, it results a more refined version of the Radon–Nikodym derivative in the realm of $\mathscr{C}^1$-differentiable manifolds. Here the Radon–Nikodym derivative is not a mere class of functions coinciding almost everywhere, but it is a true function

$$f(x) = (\mathrm{d}_x\mu)/(\mathrm{d}_xm)$$

well–defined on the complement of a null set.

5. Application to the Conditional Probability

Let us consider a differentiable map $T : X \longrightarrow Y$ between $\mathscr{C}^1$-differentiable manifolds and let P be a probability measure on X. Given a value $t \in Y$, the conditional probability P_t (interpreted intuitively as the new probability when an observer knows that the statistic T has taken the value t) *should be a probability measure on the fibre* $X_t = T^{-1}(t)$. By means of the Radon–Nikodym theorem, it is possible to define the family $\{P_t\}_{t\in Y}$ (see [1] p. 460 or [2] p. 49). Unfortunately, this family is well–defined up to coincidence almost everywhere, so that the measure P_t may be arbitrarily modified at any fixed value $t \in Y$. Therefore, for any particular choice $t \in Y$, the Radon–Nikodym theorem is unable to define P_t.

On the contrary, the differential of this yet–to–be–defined measure P_t is obvious in most usual contexts. Let us assume that T is a regular map, so that X_t is a differentiable submanifold of X and we have an exact sequence of vector spaces

$$0 \longrightarrow \mathrm{T}_x(X_t) \longrightarrow \mathrm{T}_xX \longrightarrow \mathrm{T}_tY \longrightarrow 0 \qquad (x \in X_t)\ .$$

Let us also assume that d_xP and $\mathrm{d}_tT(P)$ are Lebesgue measures. From the above exact sequence one easily obtains a Lebesgue measure μ_x on $\mathrm{T}_x(X_t)$ which is a natural candidate for the differential of P_t at x. We define the conditional probability measure, with respect to the condition $T(x) = t$, to be the unique (if it exists) absolutely continuous measure P_t on X_t such that $d_x(P_t) = \mu_x$ for any $x \in X_t$.

For example, if $T : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is defined by $T(t,x) = t$ and P has a continuous density function $f(t,x)$ with respect to the Lebesgue measure:

$$P(A) = \int_A f(t,x)\, dt\, dx$$

and $F(t) = \int f(t,x)\, dx$ is continuous, then it is easy to check that the differential of P_t at any point (t,x) must be $\big(f(t,x)/F(t)\big)dx$, where dx stands for the differential of the Lebesgue measure on $X_t = t \times \mathbb{R}$. Therefore, in this case the conditional probability P_t exists and it has the density $f(t,x)/F(t)$ with respect to the Lebesgue measure on $t \times \mathbb{R}$:

$$P_t(t \times A) = \frac{\int_A f(x,t)\, dx}{\int_{\mathbb{R}} f(t,x)\, dx}\ .$$

REFERENCES

[1] BILLINGSLEY, P., "Probability and Measure", John Wiley & Sons, New York, 1986.
[2] LEHMANN, E.L., "Testing Statistical Hypotheses", John Wiley & Sons, New York, 1986.