Notes on the Lebesgue-Radon-Nikodym Theorem

Eric A. Carlen[1]
Rutgers University

November 13, 2013

1 intorduction

1.1 DEFINITION (Mutually Singular)**.** Two positive measures μ_1 and μ_2 on a measurable space $(X, \mathcal{M})$ are *mutually singular* in case there is a measurable set A so that

$$\mu_1(A^c) = 0 \qquad \text{and} \qquad \mu_2(A) = 0 \ . \tag{1.1}$$

We denote the mutual singularity of μ_1 and μ_2 by writing $\mu_1 \perp \mu_2$.

Note that when (1.1) is satisfied, for any $E \in \mathcal{M}$,

$$\mu_1(E) = \mu_1(A \cap E) \qquad \text{and} \qquad \mu_2(E) = \mu_2(A^c \cap E) \ .$$

in this sense, "μ_1 lives on A, and μ_2 lives on the complement of A".

1.2 EXAMPLE. Let $(X, \mathcal{M}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Let μ_1 be Lebesgue measure on $\mathbb{R}$, and let μ_2 be the point mass at the origin, often called the *Dirac mass*. That is, for all $E \in \mathcal{B}_{\mathbb{R}}$, $\mu_2(E) = 1$ if $0 \in E$ and $\mu_2(E) = 0$ otherwise. Then with $A = \mathbb{R}\backslash\{0\}$, (1.1) is satisfied, and so μ_1 and μ_2 are mutually singular.

The measure μ_2 is the Lebesgue-Stieltjes measure associated to the right continuous function F where

$$F(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

through $\mu_2((a, b]) = F(b) - F(a)$.

For a second – more interesting – example, let $(X, \mathcal{M}) = ([0, 1], \mathcal{B}_{[0,1]})$, and let $F : [0, 1] \to [0, 1]$ be the Cantor function, which is continuous and monotone non-decreasing. Hence there is a unique Lebesgue-Stieltjes measure μ_2 such that $\mu_2((a, b]) = F(b) - F(a)$ for all $a < b$ in $[0, 1]$. Let C be the Cantor set. Then, as we have seen, $\mu_2(C^c) = 0$ while the Lebesgue measure of C is zero. Thus, taking μ_1 to be Lebesgue measure, and $A = C^c$, (1.1) is again satisfied, and μ_1 and μ_2 are mutually singular.

1.3 DEFINITION (Absolutely continuous)**.** Let μ_1 and μ_2 be two measures on a measurable space $(X, \mathcal{M})$. Then μ_1 is *absolutely continuous* with respect to μ_2 in case for all measurable sets A,

$$\mu_2(A) = 0 \quad \Rightarrow \quad \mu_1(A) = 0 \ . \tag{1.2}$$

1.4 EXAMPLE. Let $(X, \mathcal{M})$ be a measure space, and let μ_2 be a measure on $(X, \mathcal{M})$. Let $h \geq 0$ be an integrable function on $(X\mathcal{M}, \mu_2)$. Define a mesure μ_1 on $(X, \mathcal{M})$ by

$$\mu_1(E) = \int_E h \mathrm{d}\mu_2$$

for all $E \in \mathcal{M}$. Then, as we have seen, μ_1 is a finite measure on $(X, \mathcal{M})$ with $\mu_1(X) = \|h\|_1$.

If $\mu_2(E) = 0$, then $1_E h = 0$ a.e. with respect to μ_2, and so

$$\mu_1(E) = \int_E h \mathrm{d}\mu_2 = \int_X 1_E h \mathrm{d}\mu_2 = 0$$

since the integral of a measurable integrand that equals zero almost everywhere is zero.

The *Radon-Nikodym Theorem*, proved below, says that when μ_1 and μ_2 are finite, all examples of absolute continuity are of this type.

2 The Main Theorems

2.1 THEOREM (Lebesgue Decomposition Theorem)**.** *Let μ_1 and μ_2 be two finite measures on a measurable space $(X, \mathcal{M})$ Then there are measures $\mu_1^{(\mathrm{s})}$ and $\mu_1^{(\mathrm{ac})}$ so that*

$$\mu_1 = \mu_1^{(\mathrm{s})} + \mu_1^{(\mathrm{ac})}$$

where $\mu_1^{(\mathrm{s})}$ and μ_2 are mutually singular, and where $\mu_1^{(\mathrm{ac})}$ is absolutely continuous with respect to μ_2. Moreover, this decomposition into a singular and absolutely continuous parts is unique.

2.2 THEOREM (Radon–Nikodym Theorem)**.** *Let μ_1 and μ_2 be two finite measures on a on a measurable space $(X, \mathcal{M})$. If μ_1 is absolutely continuous with respect to μ_2, there is a function h that is integrable with respect to μ_2 such that for all $E \in \mathcal{M}$,*

$$\mu_1(E) = \int_E h \mathrm{d}\mu_1 \ , \tag{2.1}$$

and moreover, h is unique up to a.e. equivalence.

The following proof of these theorems is due to Von Neumann.

Proof. Let μ_1 and μ_2 be two finite measures on $\mathcal{M}$. Define the positive finite Borel measure ν by

$$\nu = \mu_1 + \mu_2 \ .$$

Let $\mathcal{H}$ denote $L^2(X, \mathcal{M}, \nu)$. For all $f \in \mathcal{H}$, by the fact that $\nu \geq \mu_2$, and then the Cauchy-Schwarz inequality,

$$\int_X |f| \mathrm{d}\mu_2 \leq \int_X 1|f| \mathrm{d}\nu \leq \left(\int_X 1 \mathrm{d}\nu\right)^{1/2} \left(\int_X |f|^2 \mathrm{d}\nu\right)^{1/2} = (\nu(X))^{1/2} \left(\int_X |f|^2 \mathrm{d}\nu\right)^{1/2} \ . \tag{2.2}$$

Thus, for all $f \in \mathcal{H}$, $f \in L^1(X, \mathcal{M}, \mu_2)$, and we may define a linear functional L on $\mathcal{H}$ by

$$L(f) = \int_X f \mathrm{d}\mu_2 \ .$$

It follows from (2.2) that for all $f \in \mathcal{H}$,

$$|L(f)| \le \int_X |f| \mathrm{d}\mu_2 \le (\nu(X))^{1/2} \|f\|_{\mathcal{H}} \ .$$

Therefore, L is bounded, and by the Riesz Representation Theorem, there exists a unique function $g \in \mathcal{H}$ such that

$$\int_X f \mathrm{d}\mu_2 = \int_X f g \mathrm{d}\nu \tag{2.3}$$

for all $f \in \mathcal{H}$. Since $\nu = \mu_1 + \mu_2 \ge \mu_2$, it follows immediately that for all $f \ge 0$,

$$\int_X f \mathrm{d}\nu \ge \int_X f g \mathrm{d}\nu \ge 0 \ . \tag{2.4}$$

Hence, for any $E \in \mathcal{M}$, $\nu(E) \ge \int_E g \mathrm{d}\nu \ge 0$, and this means that

$$0 \le g(x) \le 1$$

almost everywhere with respect to ν.

Now let

$$A = \{ \ x \ : \ g(x) > 0 \ \} \qquad \text{or, what is the same,} \quad A^c = \{x \ : \ g(x) = 0\} \ ,$$

and define a measure $\mu_1^{(\mathrm{s})}$ by

$$\mu_1^{(\mathrm{s})}(E) = \mu_1(A^c \cap E) \quad \text{for all} \quad E \in \mathcal{M} \ . \tag{2.5}$$

Taking $f = 1_{A^c}$ in (2.3), we see that

$$\mu_2(A^c) = 0$$

and from (2.5) that

$$\mu_1^{(\mathrm{s})}(A) = 0 \ .$$

This shows that $\mu_1^{(\mathrm{s})}$ and μ_2 are mutually singular. We next define $\mu_1^{(\mathrm{ac})}$ by

$$\mu_1^{(\mathrm{ac})} = \mu_1 - \mu_1^{(\mathrm{s})} \ ,$$

or, what is the same,

$$\mu_1^{(\mathrm{ac})}(E) = \mu_1(E \cap A)$$

for all $E \in \mathcal{M}$. It remains to find h, which we shall show is given by $h = (1-g)/g$ on A. To see this, use $\nu = \mu_1 + \mu_2$ to rewrite (2.3) as

$$\int_X f(1-g) \mathrm{d}\mu_2 = \int_X f g \mathrm{d}\mu_1 \tag{2.6}$$

for all $f \in \mathcal{H}$.

Now let E be any measurable subset of A, and for each positive integer N define

$$f_N = 1_E \min\{g^{-1}, N\} \ .$$

Since $g > 0$ on E, g^{-1} is defined and finite and

$$1_E g^{-1} = \lim_{N\to\infty} f_N \tag{2.7}$$

almost everywhere. Moreover, since f_N is bounded, it belongs to $\mathcal{H}$. Hence from (2.6),

$$\int_X f_N(1-g)\mathrm{d}\mu_2 = \int_X f_N g \mathrm{d}\mu_1 \ .$$

By (2.7) and the Lebesgue Monotone Convergence Theorem,

$$\begin{aligned} \int_E \frac{1-g}{g}\mathrm{d}\mu_2 &= \lim_{N\to\infty} \int_X f_N(1-g)\mathrm{d}\mu_2 \\ &= \lim_{N\to\infty} \int_X f_N g \mathrm{d}\mu_1 \\ &= \mu_1(E) \ . \end{aligned}$$

Taking $E = A$,

$$\int_A \frac{1-g}{g}\mathrm{d}\mu_2 = \mu_1(A) \le \mu_1(X) < \infty \ .$$

Hence the non–negative measurable function h defined by

$$h(x) = \begin{cases} 0 & \text{if } x \in A^c \\ (1-g(x))/g(x) & \text{if } x \in A \end{cases}$$

is integrable with respect to μ_2 and for all measurable sets E,

$$\mu_1^{(\mathrm{ac})}(E) = \mu_1(E \cap A) = \int_E h \mathrm{d}\mu_2 \ . \tag{2.8}$$

It follows immediately that if $\mu_2(E) = 0$, then $\mu_1^{(\mathrm{ac})}(E) = 0$, so that $\mu_1^{(\mathrm{ac})}$ is indeed absolutely continuous with respect to μ_2.

This proves the existence of the Lebesgue decomposition. As for uniqueness, suppose that

$$\mu_1 = \nu^{(\mathrm{s})} + \nu^{(\mathrm{ac})} \quad \text{and} \quad \mu_1 = \lambda^{(\mathrm{s})} + \lambda^{(\mathrm{ac})}$$

are *any* two decompositions of μ_1 into singular and absolutely continuous parts, with respect to μ_2.

Since $\nu^{(\mathrm{s})} \perp \mu_2$, there is a set $B \in \mathcal{M}$ such that $\mu_2(B^c) = 0$ and $\nu^{(\mathrm{s})}(B) = 0$. Thus, for any $E \in \mathcal{M}$,

$$\nu^{(\mathrm{ac})}(E) = \nu^{(\mathrm{ac})}(E \cap B) + \nu^{(\mathrm{ac})}(E \cap B^c) \ .$$

Since $\nu^{(\mathrm{ac})} << \mu_2$ and $\mu_2(B^c) = 0$, $\nu^{(\mathrm{ac})}(E \cap B^c) = 0$, and thus

$$\nu^{(\mathrm{ac})}(E) = \nu^{(\mathrm{ac})}(E \cap B) = \mu_1(E \cap B) - \nu^{(\mathrm{s})}(E \cap B) = \mu_1(E \cap B) \ ,$$

where the last equality is valid since $\nu^{(\mathrm{s})}(B) = 0$. Summarizing, we have shown that for all $E \in \mathcal{M}$,

$$\nu^{(\mathrm{ac})}(E) = \mu_1(E \cap B) \qquad \text{and} \qquad \nu^{(\mathrm{ac})}(B^c) = 0 \ . \tag{2.9}$$

Likwise, there is a set $A \in \mathcal{M}$ such that $\mu_2(A^c) = 0$ and $\lambda^{(\mathrm{s})}(A) = 0$. Therefore, applying the same reasoning we have applied to $\mu_1 = \nu^{(\mathrm{s})} + \nu^{(\mathrm{ac})}$, we deduce

$$\lambda^{(\mathrm{ac})}(E) = \mu_1(E \cap A) \qquad \text{and} \qquad \lambda^{(\mathrm{ac})}(A^c) = 0 \ . \tag{2.10}$$

Next, from (2.9) and (2.10), we see that

$$\mu_1(A \cap B^c) = \mu_1(A^c \cap B) = 0 \ .$$

Thus $\mu_1(A\Delta B) = 0$. Then for any $E \in \mathcal{M}$, $|\mu_1(E \cap A) - \mu_1(E \cap B)| \leq \mu_1(A\Delta B) = 0$, which means that $\mu_1^{(\mathrm{ac})}(E) = \nu^{(\mathrm{ac})}(E)$ for all $E \in \mathcal{M}$. This proves that $\lambda^{(\mathrm{ac})} = \nu^{(\mathrm{ac})}$, and hence that the Lebesgue decomposition is unique.

Finally, since for $h, \tilde{h} \in Ł^1(X, \mathcal{M}, \mu_2)$,

$$\int_E h \mathrm{d}\mu_2 = \int_E \tilde{h} \mathrm{d}\mu_2$$

for all $E \in \mathcal{M}$ if and only if $h = \tilde{h}$ a.e. with respect to μ_2. Thus, the function h in the Radon-Nikodym Theorem is unique. □

3 Transformations of Lebesgue measure under homeomorphisms with a Lipschitz inverse

Let $K \subset \mathbb{R}^n$ be compact, and let μ denote the restriction of Lebesgue measure m to K. That is, for all Borel sets E, $\mu(E) = m(E \cap K)$.

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a homeomorphism such that T^{-1} is a Lipschitz transformation on the compact set $T(K)$. That is, there exists a finite M such that

$$|T^{-1}(x) - T^{-1}(y)| \leq M|x - y|$$

for all $x, y \in T(K)$. Equivalently

$$|T(x) - T(y)| \geq \frac{1}{M}|x - y|$$

for all $x, y \in K$.

For example if T is defined on an open set U containing K , and is continuously differentiable on U, and the Jacobian determinant $\det(DT)(x)$ is non-zero everywhere on K, these conditions conditions are readily verified.

Our main goal in this section is to show that for such a transformation T, $T\#\mu$ is absolutely continuous with respect to Lebesgue measure. In a later section we shall return to the computation of the Radon-Nikodymn derivative and show that is equals $|\det(DT)(x)|$.

To prove that $T\#\mu << m$, we recall that for all Borel sets E,

$$T\#\mu(E) = m(T^{-1}(E))$$

by the very definition of $T\#\mu$. Thus $T\#\mu << m$ if and only if for all Borel sets E with $m(E) = 0$, it is the case that $m(T^{-1}(E)) = 0$. Since our hypothesis is that T^{-1} is Lipschitz on $T(K)$, it suffices to prove the following, in which we reverse the roles of T and its inverse to keep the notation simple.

3.1 THEOREM. *Let $K \subset \mathbb{R}^n$ be compact and suppose that T is a Lipschitz function on K. Let μ^* denote Lebesgue outer measure on $\mathbb{R}^n$. If $E \subset K$ is such that $\mu^*(E) = 0$, then $\mu^*(T(E)) = 0$.*

Proof. Suppose that $\mu^*(E) = 0$. Then for every $\epsilon > 0$, there exist a countable covering of E by half open rectangles R_j such that

$$\sum_{j=1}^{\infty} m(R_j) \le \epsilon \ . \tag{3.1}$$

Let $R = (a_1, b_1] \times \cdots (a_n, b_n]$ be any finite volume rectangle. We say it is *well-proportioned* in case

$$\max_{j=1,\dots,n}\{b_j - a_j\} \le 2 \min_{j=1,\dots,n}\{b_j - a_j\} \ .$$

It is easy to see that any finite volume half open rectangle can be decomposed into a finite sum of well proportioned half-open rectangles. Thus we may freely assume that all of the rectangles in (3.1) are well-proportioned.

Let R be any non-empty well proportioned half-open rectangle, and let $L = \min_{j=1,\dots,n}\{b_j - a_j\}$. Then

$$m(R) \ge L^n \qquad \text{and} \qquad \text{diam}(R) \le \sqrt{n}2L \ ,$$

and hence

$$\text{diam}(R) \le \sqrt{n}2(m(R))^{1/n} \ .$$

Let M be the Lipschitz constant of T, so that $|T(x) - T(y)| \le M|x - y||$ for all $x, y \in K$ It follows that

$$\text{diam}(T(R)) \le M\sqrt{n}2(m(R))^{1/n} \ ,$$

and then that for any $\eta > 0$, $T(R)$ is contained in a half open rectangle that is a cube of side length $(1 + \eta)M\sqrt{n}2(m(R))^{1/n}$. Call this rectangle $\widetilde{R}$, and note that

$$m(\widetilde{R}) \le ((1 + \eta)M\sqrt{n}2)^n(m(R)) \ .$$

Now going back to our countable covering $\{R_j\}$ of E by well-proportioned rectangles, we see that $\{\widetilde{R}_j\}$ is a countable covering of $T(E)$ by half-open rectangles, and

$$\sum_{j=1}^{\infty} m(\widetilde{R}_j) \le ((1 + \eta)^n M\sqrt{n}2)^n \sum_{j=1}^{\infty} m(R_j) \le ((1 + \eta)M\sqrt{n}2)^n \epsilon \ .$$

Since $\epsilon > 0$ is arbitrary, it follows that $\mu^*(T(E)) = 0$.

□