On Radon Measure

David Kent & Lindsay Mercer

12/3/2016

1 Introduction

In this paper we define and discuss Radon measures, which are measures on certain topological spaces which interact nicely with the underlying topology. Radon measures allow for approximation by open sets or compact sets. They also allow for a unique representation of linear functionals as integrals so that measure and integration can be handled with the tools of functional analysis [9]. We will also address Radon spaces, on which every finite measure is Radon, and metrics on a space of Radon measures. In particular, the Wasserstein distance defines a metric on Radon probability measures which has close ties to transport optimization and other applied problems.

In general, this paper summarizes and follows the treatment of the subject in Folland's *Real Analysis* [2], and will reference other sources when used. Many propositions and theorems will be supplied without proof, and the reader is referred to [2] for the details.

2 Definition and properties of Radon measure

A topological space X is called **locally compact** if every point in X has a compact neighborhood. X is called a **Hausdorff space** if every pair of distinct points in X have disjoint neighborhoods. A useful consequence of this is that compact sets are closed in Hausdorff spaces [8]. For the rest of our discussion of Radon measures, X will refer to a locally compact Hausdorff (LCH) space, and we will define measures on the σ-algebra $\mathcal{B}$, the Borel sets of X.

Definition 2.1 (Regularity)**.** A measure μ is called **outer regular** on Borel-measurable E if

$$\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ open}\}$$

and μ is called **inner regular** on Borel-measurable E if

$$\mu(E) = \sup\{\mu(K) : K \subset E, K \text{ compact}\}.$$

μ is called **regular** if it is both inner and outer regular on all Borel sets.

Essentially, these regularity conditions allow the measure of a set to be approximated by the measures of open sets from the outside or compact (and hence closed) sets from the inside. Using these definitions, we are ready to define Radon measure on an LCH space X.

Definition 2.2 (Radon measure)**.** A **Radon measure** is a Borel measure that is finite on compact sets, outer regular on all Borel sets, and inner regular on open sets.

Note that some authors define a Radon measure μ on the Borel σ-algebra of any Hausdorff space to be any Borel measure that is inner regular on open sets and **locally finite**, meaning that for every point $x \in X$ there is an open neighborhood of x with finite measure [9, 7]. On an LCH space these definitions are equivalent and some of the following results, including our formulation of the Riesz Representation theorem, are dependent on the locally compact condition, so we use the LCH space definition given here.

Radon measures on LCH spaces have several useful properties. Recall that a measure μ is **σ-finite** if the space X can be written as a countable union of finite-measure sets. Similarly, for the corollary below, we say a set or a space X is **σ-compact** if it is a countable union of compact sets. These results state that when working with σ-finite sets or measures we gain regularity.

Proposition 2.1. *A Radon measure is inner regular on all of its σ-finite sets*

Corollary 2.1.1.

1. *If a Radon measure is σ-finite then it is regular.*
2. *If X is σ-compact, every Radon measure on X is regular.*

The proof of 2.1 involves using the inner and outer regularity properties of the Radon measure to approximate the measure of σ-finite sets of finite measure, and for sets of infinite measure to apply that technique to a sequence of sets, each of which has finite measure. From this, 2.1.1 follows directly from the definitions of σ-finite and σ-compact.

Proposition 2.2. *Let $(X, \mathcal{B}, \mu)$ be a σ-finite Radon measure space, and let $E \in \mathcal{B}$. Then*

1. *for any $\varepsilon > 0$ there exist U open and F closed such that $F \subset E \subset U$ and $\mu(U \setminus E) < \varepsilon$.*
2. *there exist a set $A := \bigcup_{n=1}^{\infty} F_n$, where each F_n is closed, and a set $B := \bigcap_{n=1}^{\infty} G_n$, where each G_n is open ($n \in \mathbb{N}$), such that $A \subset E \subset G$ and $\mu(B \setminus A) = 0$.*

Proof. (Sketch) Consider a partition of E into disjoint subsets of finite measure, and approximate each subset with an open set to within a factor of ε. Let U be the countable union of each of those open sets such that $\mu(U \setminus E) < \varepsilon/2$. Similarly construct an open set V which is the countable union of open approximations of a partition of E^c with $\mu(V \setminus E^c) < \varepsilon/2$. Then let $F = V^c$, and find that

$$\mu(U \setminus F) = \mu(U \setminus E) + \mu(E \setminus F) = \mu(U \setminus E) + \mu(V \setminus E^c) < \varepsilon.$$

2.) follows from 1.) in a straightforward way, from taking the limit over the countable unions so that in the difference $\mu(B \setminus A) \to 0$. □

Finally, on a different note, we present a useful approximation result for L^p spaces under Radon measures.

Proposition 2.3. *If μ is a Radon measure on X, $C_c(X)$ is dense in $L^p(\mu)$ for $1 \leq p < \infty$.*

This last result is shown by first restricting our attention to sets $E \in \mathcal{B}$ of finite measure (since if $\mu(E) = \infty$, then $\chi_E \notin L^p(\mu)$). We then show that in the L^p norm we can approximate the characteristic function χ_E with functions in $C_c(X)$. This is done by applying 2.2 to obtain a compact $K \subset E$ and an open $U \supset E$ such that $\mu(U \setminus K) < \varepsilon$. Using Urysohn's Lemma (3.1 below) we can find an $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_U$, employ our ε-bound on the difference, and thus approximate any L^p simple function. These are in turn dense in L^p and so we have the proposition.

3 Riesz Representation Theorem (take 1)

In order to prove a powerful representation result relating linear functionals and Radon measures, we will first need to establish an important topological lemma and some definitions.

Theorem 3.1 (Urysohn's Lemma)**.** *If X is a locally compact Hausdorff space and $K \subset U \subset X$, where K is compact and U is open, there exists a continuous $f : X \to [0, 1]$ such that $f = 1$ on K and $f = 0$ outside a compact subset of U.*

Urysohn's lemma is very helpful in a situation which crops up often in these regularity proofs; we have a compact K contained in an open U, and we are trying to approximate a characteristic function with a continuous function. Urysohn's lemma says that such a continuous function exists, approximating the characteristic functions of K and U up to $U \setminus K$.

When considering any real or complex-valued function f on the domain of X, we will define the **support of** f, denoted $\text{supp}(f)$, as the closure of the set $\{x \in X : f(x) \neq 0\}$. Going forward, we shall be concerned chiefly with those functions which have **compact support**, meaning that $\text{supp}(f)$ is a compact set. We then define $C_c(X)$ as the space of all continuous functions $f : X \to \mathbb{C}$ with compact support. Note that since Radon measures are finite on compact sets, the support of any $f \in C_c(X)$ has finite measure under any Radon measure and the image of each $f \in C_c(X)$ is bounded.

Let I be a real-valued linear functional on $C_c(X)$, that is, for $f \in C_c(X)$, $I : f(x) \to \mathbb{R}$. Then we say I is **positive** if $I(f) \geq 0$ whenever $f \geq 0$. In the following proposition, we formally state that positive linear functionals are bounded on compact sets. In it, we use the **uniform norm** of f, which is equivalent on $C_c(X)$ to the infinity norm of f, defined as $\|f\|_u := \sup\{|f(x)| : x \in X\}$.

Proposition 3.2. *If I is a positive linear functional on $C_c(X)$, for each compact set $K \subset X$ there exists a constant C_K such that $|I(f)| \leq C_K \|f\|_u$ for all $f \in C_c(X)$ such that* $\text{supp}(f) \subset K$.

Now let μ be a Borel measure on X, let $\mathcal{B}$ be the Borel σ-algebra on X, and let $E \in \mathcal{B}$. When μ is finite on compact sets, then $\int f d\mu$ is finite for all $f \in C_c(X)$ and $\int f d\mu \geq 0$ whenever $f \geq 0$. Hence, the mapping $I : f \mapsto \int f$ is a positive linear functional. In particular, since Radon measures are Borel and finite on compact sets, the integral with respect to any Radon measure is a positive linear functional on $C_c(X)$. The Riesz Representation Theorem asserts that this is in fact the only kind of positive linear functional on $C_c(X)$: any positive linear functional can be expressed as the integral with respect to a Radon measure.

In order to provide some intuition for how we might find such a measure μ relating a positive linear functional I to the integral with respect to μ, observe that for any measure μ, $\mu(E) = \int \chi_E d\mu$. Since we want I and the integral to match, we might consider setting $\mu(E) = I(\chi_E)$ for every $E \in \mathcal{B}$, in which case we would have $I(\chi_E) = \int \chi_E d\mu$ for every $E \in \mathcal{B}$, which seems like a good start. However, characteristic functions are neither continuous nor have compact support in general, and since I is a linear functional on $C_c(X)$, we cannot do the above exactly.

Instead, given an open set U, we will consider the supremum of $I(f)$ for $0 \leq f \leq \chi_U$. We will also need to require that $\text{supp}(f) \subset U$. We will write $f \prec U$ in the case that $0 \leq f \leq \chi_U$ and $\text{supp}(f) \subset U$. This brings us to the full statement of the Riesz Representation Theorem.

Theorem 3.3 (The Riesz Representation Theorem)**.** *If I is a positive linear functional on $C_c(X)$, there is a unique Radon measure μ on X such that $I(f) = \int f d\mu$ for all $f \in C_c(X)$. Moreover, μ satisfies*

$$\mu(U) = \sup\{I(f) : f \in C_c(X) \text{ and } f \prec U\} \text{ for all open } U \subset X, \tag{1}$$

$$\mu(K) = \inf\{I(f) : f \in C_c(X) \text{ and } f \geq \chi_K\} \text{ for all compact } K \subseteq X. \tag{2}$$

Proof. (Sketch) Proving the uniqueness of the measure μ given by this theorem is relatively straightforward. Using Urysohn's lemma and the inner regularity of μ on open sets, we first demonstrate that 3.3(1) is satisfied. It is then apparent that μ is determined by I on open sets, and because of outer regularity, is determined by I on Borel sets, and thus is unique.

Proving existence is much more involved, but still straightforward. Briefly, (i) an outer measure is constructed on X, (ii) open sets are shown to be outer-measurable and a Borel measure is constructed from the outer measure, which (iii) is shown to satisfy 3.3(2) above. It is demonstrated that μ is inner regular and hence Radon using that property. Finally, (iv) it is shown that $I(f) = \int f d\mu$ for all $f \in C_c(X)$.

For (i), we define $\mu(U) = \sup\{I(f) : f \in C_c(X) \text{ and } f \prec U\}$ for open $U \subset X$. Then, for arbitrary $E \subset X$, we let $\mu^*(E) = \inf\{\mu(U) : E \subset U, U \text{ open}\}$. In order to show that μ^* is an outer measure, it suffices to show that if (U_i) is a sequence of open sets with $\bigcup U_i = U$, then $\mu(U) \leq \sum \mu(U_i)$, from which it follows that $\mu^*(E) = \inf \{\sum \mu(U_i) : U_i \text{ open }, E \subset \bigcup U_i\}$, which defines an outer measure.

The proof of (ii) teases out the definitions of μ on open sets and μ^* on arbitrary sets to show that open sets are μ^*-measurable. We prove (iii) from (ii), since it follows from Carathéodory's extension theorem that all Borel sets are μ^*-measurable and that $\mu = \mu^*|\mathcal{B}_X$ is a measure. This measure μ satisfies 3.3(1) and is outer regular by definition. From (iii), it follows that μ is inner regular.

To prove (iv), it suffices to show that $I(f) = \int f d\mu$ for all $f \in C_c(X, [0,1])$, since $C_c(X)$ is the linear span of $C_c(X, [0,1])$. This is done by constructing a sequence (f_i) of N functions which sum to f, and using the characteristic function of the support of each f_i, 3.3(2), and outer regularity to demonstrate that $|I(f) - \int f d\mu| \leq \frac{1}{N}\mu(\mathrm{supp}(f))$. Since $\mu(\mathrm{supp}(f))$ is finite and N is arbitrary, $I(f) = \int f d\mu$. □

4 Radon space

We will now consider the class of spaces where every finite measure is Radon [6].

Definition 4.1. A Hausdorff space X is called a **Radon space** if every finite Borel measure on X is a Radon measure.

In Folland, it is shown that every second countable space is a Radon space [2]. In fact, the theorem given is slightly stronger:

Theorem 4.1. *Let X be a locally compact Hausdorff space in which every open set is σ-compact (which is the case, for example, if X is second countable). Then every Borel measure on X which is finite on compact sets is regular and hence Radon.*

Proof. (Sketch) Consider a Borel measure μ which is finite on compact sets. Then each $f \in C_c(X)$ has finite integral with respect to μ, so $I(f) = \int f d\mu$ is a positive linear functional on $C_c(X)$. We can then apply the Riesz Representation Theorem 3.3 to obtain a Radon measure ν associated with I. First we show that $\mu(U) = \nu(U)$ for all open U. Using the σ-compactness of U, choose a sequence (K_n) of compact K such that $\bigcup K_i = U$. Using Urysohn's Lemma 3.1, we can construct a sequence (f_n) of functions with compact support, increasing pointwise to χ_U. Then we get $\mu(U) = \lim \int f_n d\mu = \lim \int f_n d\nu = \nu(U)$, the last inequality following from the monotone convergence theorem.

Next we show regularity. Let E be a Borel set, $\varepsilon > 0$. Then by 2.2, there are closed F and open V such that $F \subset E \subset V$, with $\nu(V \setminus F) < \varepsilon$. Since $V \setminus F$ is open, we get that $\mu(V \setminus F) < \varepsilon$. By subadditivity and monotonicity of measures, this gives us $\mu(V) \leq \mu(E) + \varepsilon$, so μ is outer regular. Similarly, we have $\mu(F) \geq \mu(E) - \varepsilon$. Since F is closed and E σ-compact, F is σ-compact, so we can choose a increasing sequence (K_i) of compact $K_i \subset F$ with $\bigcup K_i = F$, in which case $\mu(K_i)$ converges to $\mu(F)$, so μ is inner regular. □

Thus every finite measure on a second countable space is Radon. Since every separable metric space is second countable, every separable metric space is a Radon space.

5 Riesz Representation Theorem (take 2)

We say that a function f **vanishes at infinity** if for any $\varepsilon > 0$, $\{x \in X : |f(x) \geq \varepsilon\}$ is compact.

We will now turn our attention towards $C_0(X)$, the spaces of continuous functions which vanish at infinity. In any locally compact Hausdorff space, $C_0(X)$ is the uniform closure of $C_c(X)$. In this section, we will give a complete description of $C_0(X)^*$, the bounded linear functionals on $C_0(X)$. The first important fact is that real linear functionals on $C_0(X)$ have a "Jordan decomposition."

Theorem 5.1. *If $I \in C_0(X, \mathbb{R})^*$, there is are positive linear functionals $I^+, I^- \in C_0(X, \mathbb{R})^*$ such that $I = I^+ - I^-$.*

A **signed Radon measure** is a signed Borel measure whose positive and negative variations are Radon measures. We will denote the space of signed Radon measures on X by $M(X)$ and for $\mu \in M(X)$ define $\|\mu\| = |\mu|(X)$. It can be shown that $M(X)$ is a complete normed vector space.

Theorem 5.2 (The Riesz Representation Theorem)**.** *Let X be a locally compact Hausdorff space, and for $\mu \in M(X)$ and $f \in C_0(X)$, let $I_\mu(f) = \int f d\mu$. Then the map $\mu \to I_\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.*

Proof. (Sketch) To see that there is a bijection from $M(X)$ to $C_0(X)^*$, we first note that every bounded linear functional on $C_c(X)$ extends continuously to $C_0(X)$, since $C_0(X)$ is the uniform closure of $C_c(X)$. So each linear functional is given by integration against some Radon measure. On the other hand, $|\int f d\mu| \leq \int |f| d|\mu| \leq \|f\|_\mu \|\mu\|$, so every integral against a Radon measure is a bounded linear functional.

In order to show that the norms are equal, first note that the above inequality gives that $\|I_\mu\| \leq \|\mu\|$. To show the other direction, we approximate the Radon-Nikodym derivative $\frac{d\mu}{d|\mu|}$ by some $f \in C_c(X)$ such that $\|\mu\| \leq |\int f d\mu| + \varepsilon$, from which it follows that $\|\mu\| \leq \|I_\mu\|$. □

The Riesz Representation Theorem gives a complete characterization of the space of bounded linear functionals on $C_0(X)$ - it is exactly the space of signed Radon measures.

6 Radon and Wasserstein metrics

Let $M_+(X)$ be the space of finite positive Radon measures. $M_+(X) \subset M(X)$ is clearly closed under addition and scalar multiplication for scalars $c \geq 0$ and contains $\mu \equiv 0$, and thus is a pointed convex cone. We now define a distance on this space.

Definition 6.1 (Radon distance)**.** The **Radon distance** between measures $\mu, \nu \in M_+(X)$ is

$$\rho(\mu, \nu) := \sup \left\{ \int_X f(x) d(\mu - \nu) : \text{continuous real-valued } f : X \to [-1, 1] \right\} \tag{3}$$

Since all elements of $M(X)$ are finite measures, we are guaranteed this supremum exists in $\mathbb{R}$. It can be shown that ρ is a metric on $M_+(X)$ and that $M_+(X)$ is a complete metric space under ρ [9].

The Radon metric may seem to be a natural choice for applications on $M_+(X)$, but in fact it is often too restrictive. For example, let us consider the oft-used space of **probability measures** in $M_+(X)$, defined:

$$\mathcal{P}(X) := \{\mu \in M_+(X) : \mu(X) = 1\}.$$

Under the Radon metric, $\mathcal{P}(X)$ is not sequentially compact, meaning there is no guarantee that every sequence has a convergent subsequence [9]. For this and other reasons, other more useful metrics have been defined on $\mathcal{P}(X)$, chief among them, the Wasserstein metric.

First, we need to define $\Pi(\mu, \nu)$ as the set of all **couplings** in $\mathcal{P}(X \times X)$: pairings of measures on $(X, \mathcal{B})$ which preserve the marginal distributions of μ and ν respectively on X. That is,

$$\Pi(\mu, \nu) := \left\{ \pi \in \mathcal{P}(X \times X) : \begin{array}{l} \pi(A \times X) = \mu(A) \\ \pi(X \times A) = \nu(A) \end{array}, \text{ for each } A \in \mathcal{B} \right\}$$

For a Radon space X imbued with a metric d, let us define $\mathcal{P}_p(X)$ as the collection of probability measures with a **finite p-th moment**, for $p \in [1, \infty)$. This means that for any $\mu \in \mathcal{P}_p(X)$, around any fixed central point $x_0 \in X$, $\int_X d(x, x_0)^p d\mu(x) < \infty$.

Definition 6.2 (p-th Wasserstein distance)**.** For $\mu, \nu \in \mathcal{P}_p(X)$ the **p-th Wasserstein distance** between μ and ν is defined:

$$W_p(\mu, \nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y)^p d\pi(x, y) \right)^{1/p} \tag{4}$$

or equivalently:

$$\inf_{\pi \in \Pi(\mu,\nu)} \|d\|_{L^p(\pi)}$$

It can be shown that if X is a compact metric space, i.e. if X is complete and totally bounded under metric d, then W_p is a metric on $\mathcal{P}_p(X)$, sometimes called the **Kantorovich metric** [1, 9, 5].

The **support of a measure** μ is the largest closed subset of X for which every open neighborhood of any point in the set has strictly positive measure. When μ and ν have bounded support, which is certainly the case for X compact, the W_1 metric has a dual formulation which is closely related to the Radon metric [1, 4]

Theorem 6.1 (Kantorovich & Rubinstein (1958): case $p = 1$)**.**

$$W_1(\mu, \nu) = \sup \left\{ \int_X f(x) d(\mu - \nu) : \text{ continuous } f : X \to \mathbb{R}, \mathrm{Lip}(f) \leq 1 \right\} \tag{5}$$

on $\mathcal{P}(X)$*, where* $\mathrm{Lip}(f)$ *is the minimal Lipschitz constant of* f*.*

The formulations (5) and (3) of the Wasserstein-1 and Radon metrics are quite similar. In fact, it is easily shown (see [3] for proof) that, if C is the constant bound for d on X compact, then

$$2\, W_1(\mu, \nu) \leq C\, \rho(mu, \nu).$$

From this, we infer that convergence in the Radon metric implies convergence in the Wasserstein-1 metric [10], although the converse is not necessarily true.

In the field of tranportation, the Wasserstein metric is closely tied to both the Monge and Kantorovich transport problems, which are concerned with optimally moving mass between two probability distributions on the same space. In the non-linear Monge problem, we seek a map between two probability spaces which will minimize the integral of a cost function on the domain space. By contrast, the linear Kantorovich problem seeks a probability coupling π which minimizes the integral of a cost function on the product space. This eventually lead to the Wasserstein metric (thus sometimes called the Kantorovich-Rubinstein metric) and the dual formulation (6.1) [1]. W_1 is also referred to as the *"earth mover's distance"*, at least in part because Monge's original problem was constructed in the context of optimally transporting construction soil, where the probability distributions represented the conserved volume of earth, and the cost function represented distance moved.

Many other applications of the Wasserstein distance exist, including but certainly not limited to: [4]

- measuring the rate of convergence of probability measures
- measuring geodisic distances
- coupling stochastic differential equations
- heat distribution as gradient flows

References

[1] Vladimir I. Bogachev and Aleksandr V. Kolesnikov. The Monge-Kantorovich problem: achievements, connections, and perspectives. *Russian Mathematical Surveys*, 67(5):785–890, November 2012.

[2] Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*, chapter 7, pages 211–233. John Wiley & Sons, 2nd edition, May 1999.

[3] Alison L. Gibbs and Francis Edward Su. On choosing and bounding probability metrics. *International Statistical Review*, 70(3):419–435, 2002.

[4] Kazumasa Kuwada. Probability analysis on Wasserstein distance. In *Probabilistic symposium - Research meeting*, December 19-22, 2011. www.math.tohoku.ac.jp/~aida/workshop/H23-Daisympo/kuwada.pdf.

[5] Filippo Santambrogio. The Wasserstein Distances. Course notes, URL: https://www.math.u-psud.fr/~filippo/Wp.pdf, March 20, 2011.

[6] Laurent Schwartz. *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. Oxford University Press, 1973.

[7] William (http://math.stackexchange.com/users/327486/william). What is actually the standard definition for Radon measure? Mathematics Stack Exchange. URL: http://math.stackexchange.com/q/1858742 (version: 2016-07-13).

[8] Wikipedia. Hausdorff space — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Hausdorff_space, 2016. [Online; accessed 27-November-2016].

[9] Wikipedia. Radon measure — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Radon_measure, 2016. [Online; accessed 27-November-2016].

[10] Wikipedia. Wasserstein metric — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Wasserstein_metric, 2016. [Online; accessed 27-November-2016].