MATH 650. THE RADON-NIKODYM THEOREM

This note presents two important theorems in Measure Theory, the Lebesgue Decomposition and Radon-Nikodym Theorem. They are not treated in the textbook.

1. Closed subspaces

Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

Definition 1. A *subspace* of H is a subset of H such that, if it contains x and y, then it also contains $\alpha x + \beta y$, for every pair of complex numbers α and β.

A *closed subspace* is a subspace which is also a closed subset of H, that is, every Cauchy sequence in the subspace converges to a vector in the subspace.

Exercise 1. Let $H = \mathcal{L}^2(X, \mu)$. Show that the collection of simple functions is a subspace of H, but not closed in general.

Example 1. Let $H = \mathcal{L}^2(X, \mu)$, where $X = [0, 1]$ and μ is Lebesgue measure. The following are examples of subspaces of H.

(1) the set of all functions $f \in H$ such that $f(x) = f(-x)$ for almost every $x \in X$.
(2) the set of all functions f such that $\int f \cdot \mu = 0$.
(3) the set of all functions f which are bounded on a subset of X of measure $\mu(X)$.

The first two examples are closed subspaces, but the last one is not.

Example 2. Let H be the set of sequences of complex numbers $\{z_n\}$ such that $\sum_{n=1}^{\infty} |z_n|^2 < \infty$. Then the subset of H consisting of sequences $\{z_n\}$ with only finitely many non-zero terms is a subspace but not a closed subspace.

2. Orthogonal Decomposition

A vector x is *orthogonal* to a subset A of H if the inner product $\langle x, a \rangle = 0$ for every $a \in A$. The set of all vectors orthogonal to A is denoted by $A^{\perp}$.

Exercise 2. Let A be a subset of H.

(1) Show that $A^{\perp}$ is a closed subspace.

(2) If A is a closed subspace, then show that $(A^{\perp})^{\perp} = A$.

Proposition 1. *If K is a closed subspace of H, if x is a vector and if $d = \inf\{\|y - x\| \mid y \in K\}$, then there exists a unique vector $y_0 \in K$ such that $\|y_0 - x\| = d$.*

Proof. Let y_n be a sequence of vectors in K such that $\|x - y_n\| \to d$ as $n \to \infty$. It follows from the parallelogram law (Exercise 3, §3.2) applied to the vectors $x - y_n$ and $x - y_m$ that

$$\|y_n - y_m\|^2 = 2\|x - y_n\| + 2\|x - y_m\|^2 - 4\|x - (y_n + y_m)/2\|^2$$

for every n and m. Since $(y_n + y_m)/2 \in K$, it follows that

$$\|x - (y_n + y_m)/2\|^2 \geq d^2$$

and hence that

$$\|y_n - y_m\|^2 \leq 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2.$$

As $n \to \infty$ and $m \to \infty$, the right side of this inequality tends to $2d^2 + 2d^2 - 4d^2 = 0$, so that $\{y_n\}$ is a Cauchy sequence, and so it converges in H. If $y_n \to y_0$, then $y_0 \in K$ (because K is a closed subspace and $y_n \in K$) and, by the continuity of the norm (Theorem 14, §3.2),

$$\|y_0 - x\| = \lim_n \|y_n - x\| = d.$$

If y_1 is another vector in K such that $\|y_1 - x\| = d$, then $(y_1 + y_0)/2 \in K$, so the definition of d implies $\|x - (y_0 + y_1)/2\|^2 \geq d^2$. This and the parallelogram law give

$$\begin{aligned} \|y_1 - y_0\|^2 &= 2\|x - y_0\|^2 + 2\|x - y_1\| - 4\|x - (y_0 + y_1)/2\|^2 \\ &\leq 2d^2 + 2d^2 - 4d^2 \end{aligned}$$

Thus $\|y_0 - y_1\| = 0$ and so $y_0 = y_1$ by the properties of the norm. □

Theorem 1 (Othogonal Decomposition)**.** *Let $K \subset H$ be a closed subspace. Then every vector x can be written in a unique way as a sum $x = Tx + Px$, where $Tx \in K$ and $Px \perp K$.*

Proof. Let $Tx \in K$ be the vector obtained in Proposition 1, and let $Px = x - Tx$. It has to be shown that $Px \perp K$, that is, that $\langle x - Tx, y\rangle = 0$ for all y in K. Let $y \in K$ with $\|y\| = 1$. For every complex number α, the vector $Tx + \alpha y \in K$ because K is a subspace. Therefore, for all α,

$$\|x - Tx\|^2 \leq \|x - Tx - \alpha y\|^2$$

by the minimizing property of Tx. This inequality simplifies to

$$0 \leq |\alpha|^2 - \overline{\alpha}\langle x - Tx, y\rangle - \alpha\langle y, x - Tx\rangle.$$

Taking in particular $\alpha = \langle x - Tx, y\rangle$, it obtains that $0 \le -|\langle x - Tx, y\rangle|^2$, hence that $x - Tx \perp K$.

Uniqueness of the orthogonal decomposition is an easy exercise. $\square$

Corollary 1. *Let K be a closed subspace such that $K \neq H$. Then there exists a non-zero vector $x \in K^{\perp}$.*

3. Riesz representation theorem

A *linear functional* on H is a map $f : H \to \mathbf{C}$ such that

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

for every pair of vectors x and y and every pair of complex numbers α and β.

Definition 2. A *continuous linear functional* on H is a linear functional $f : H \to \mathbf{C}$ which is continuous, that is, whenever $x_n \to x$ in H, then $f(x_n) \to f(x)$ in $\mathbf{C}$.

Exercise 3. Let $f : H \to \mathbf{C}$ be a linear functional and suppose that there exists a constant $C > 0$ such that

$$|f(x)| \le C\|x\|$$

for every $x \in H$. Show that f is continuous.

Example 3. Let $H = \mathcal{L}^2(X, \mu)$. Then

$$f \mapsto \int_X f d\mu$$

is a linear functional. If the measure $\mu(X) < \infty$, then it is also a continuous linear functional. Indeed

$$\left| \int_X f \cdot \mu \right| \le \mu(X)\|f\|_2,$$

by the Schwarz inequality (Theorem 7, §3.2).

More generally, if $z \in H$, then

$$x \in H \mapsto \langle x, z\rangle$$

is a continuous linear functional on H. It turns out that every continuous linear functional is of this form.

Theorem 2 (Riesz Represenation)**.** *If f is a continuous linear functional on H, then there exists a unique vector $z \in H$ such that*

$$f(x) = \langle x, z\rangle$$

for every $x \in H$.

Proof. If $f(x) = 0$ for every $x \in H$, then take $z = 0$. Assume thus f is not identically 0. Let $K = \{x \in H \mid f(x) = 0\}$. Then K is a subspace of H, and is also closed because f is continuous and $K = f^{-1}(0)$.

Because $f \not\equiv 0$, the closed subspace $K \neq H$. Thus, by Corollary 1, there exits a non-zero vector $y \in H$ such that $y \perp K$. It may be assumed that $\|y\| = 1$ (otherwise replace y by $y/\|y\|$).

Put $u = (f(x))y - (f(y))x$. Then $u \in K$ because $f(u) = f(x)f(y) - f(y)f(x) = 0$. Therefore $\langle u, y\rangle = 0$. But

$$\langle u, y\rangle = \langle (f(x))y - (f(y))x, y\rangle = f(x) - f(y)\langle x, y\rangle$$

so that

$$f(x) = f(y)\langle x, y\rangle.$$

Hence, $f(x) = \langle x, z\rangle$ with $z = \overline{f(y)}y$.

To see that z is unique, note that if $\langle x, z\rangle = \langle x, z'\rangle$ for all $x \in H$, then $u = z - z'$ is such that $\langle x, u\rangle = 0$ for all $x \in H$, hence $u = 0$. $\square$

4. The Lebesgue Decomposition Theorem

Let $(X, \mathcal{F})$ be a measure space, and let μ and ν be two measures on X.

Definition 3. The measure ν is said to be *absolutely continuous with respect to* μ, in symbols $\nu \prec \mu$, if $\nu(E) = 0$ whenever $\mu(E) = 0$.

Example 4. If f is a non-negative μ-integrable function on X, then $\nu(E) = \int_E f \cdot \mu$ defines a measure on X which is absolutely continuous with respect to μ.

Exercise 4. Suppose that μ and ν are finite measures on $(X, \mathcal{F})$ Then ν is absolutely continuous with respect to μ if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\nu(E) < \varepsilon$ for every $E \in \mathcal{F}$ such that $\mu(E) < \delta$.

Definition 4. The measures μ and ν are *singular*, written $\mu \perp \nu$, if there is an element $E \in \mathcal{F}$ such that $\mu(E) = 0 = \nu(X \setminus E)$.

Example 5. Let $X = \mathbf{R}$, $\mathcal{F}$ the Borel sets. Let μ be Lebesgue measure and ν be the measure given by $\nu(E) = 1$ if $0 \in E$ and $\nu(E) = 0$ if $0 \notin E$. Then μ and ν are singular.

Theorem 3 (Lebesgue Decomposition)**.** *Let $(X, \mathcal{F})$ be a measurable space, and let μ and ν be two finite measures on X. Then there exist unique measures ν_a and ν_s such that $\nu = \nu_a + \nu_s$, $\nu_a \prec \mu$ and $\nu_s \perp \mu$.*

Proof. Let H be the Hilbert space $H = \mathcal{L}^2(X, \nu + \mu)$. Since the measures ν and μ are finite, so is $\nu + \mu$. Moreover, if $f \in H$, then also $f \in \mathcal{L}^2(X, \nu)$. Therefore

$$f \in H \mapsto \int_X f \cdot \nu$$

is a continuous linear functional on H, because $\nu \leq \mu + \nu$ and Example 3. By Theorem 2, there exists $g \in H$ such that

$$(\dagger) \qquad \int f \cdot \nu = \int fg \cdot (\nu + \mu),$$

for all $f \in H$. Then

$$(*) \qquad \int f(1-g) \cdot (\mu + \nu) = \int f \cdot \mu.$$

This identity implies that g is real valued. Furthermore, $g \geq 0$, for otherwise, take f to be the characteristic function of the set $\{g(x) < 0\}$ for a contradiction. Likewise, $g \leq 0$. Thus it may be assumed that $0 \leq g(x) \leq 1$ for all x. The monotone convergence theorem implies that $(*)$ holds for all $f \geq 0$.

Let $A = \{g = 1\}$ and $B = X \setminus A$. Then letting $f = \chi_A$ in $(*)$ gives $\mu(A) = 0$. For $E \in \mathfrak{F}$, set

$$\nu_s(E) := \nu(E \cap A), \qquad \text{and} \qquad \nu_a(E) := \nu(E \cap B)$$

Then ν_s and ν_a are measures, $\nu = \nu_s + \nu_a$, and $\nu_s \perp \mu$.

If $E \in \mathfrak{F}$ and $\mu(E) = 0$ and $E \subset B$, then $\int_E (1-g)\cdot(\mu+\nu) = \int_E \mu = 0$ by $(*)$ with $(1-g) > 0$ on E, so $(\mu+\nu)(E) = 0$ and $\nu(E) = \nu_a(E) = 0$. Hence $\nu_a \prec \mu$, and the existence part of the Lebesgue Decomposition is proved.

To prove uniqueness, suppose that there is another decomposition $\nu = \rho + \sigma$, with $\rho \prec \mu$ and $\sigma \perp \mu$. Then $\rho(A) = 0$ because $\mu(A) = 0$. Thus for all $E \in \mathfrak{F}$

$$\nu_s(E) = \nu(E \cap A) = \sigma(E \cap A) \leq \sigma(E)$$

Thus $\nu_s \leq \sigma$ and $\rho \leq \nu_a$. Then $\sigma - \nu_s = \nu_a - \rho$ is a measure both absolutely continuous and singular with respect to μ, so it is 0. Hence $\rho = \nu_a$ and $\sigma = \nu_s$. □

Example 6. Let $X = \mathbf{R}$, let μ be Lebesgue measure on $[0, 2]$ and ν be Lebesgue measure on $[1, 3]$. Then ν_a is Lebesgue measure on $[1, 2]$ and ν_s is Lebesgue measure on $[2, 3]$.

5. The Radon-Nikodym Theorem

Theorem 4 (Radon-Nikodym)**.** *Let $(X, \mathcal{F}, \mu)$ be a finite measure space. If ν is a finite measure on $(X, \mathcal{F})$, absolutely continuous with respect to μ, then there exists an integrable function $h \in \mathcal{L}^1(X, \mu)$ such that*

$$\nu(E) = \int_E h \cdot \mu$$

for all $E \in \mathcal{F}$. Any two such h are equal almost everywhere with respect to μ.

Proof. Note that the Lebesgue decomposition theorem gives $\nu = \nu_a + \nu_s$. Since $\nu \prec \mu$, it must be that $\nu = \nu_a$ and so $\nu_s = 0$. Let $h = g/(1-g)$ on B and $h \equiv 0$ on A. If $E \in \mathcal{F}$, let $f = h\chi_E$ in $(*)$. Then, by $(*)$ and $(\dagger)$,

$$\int_E h \cdot \mu = \int_{B \cap E} g \cdot (\mu + \nu) = \nu(B \cap E) = \nu(E).$$

To prove uniqueness, let k be another μ-integrable function such that $\nu(E) = \int_E k \cdot \mu$, for all $E \in \mathcal{F}$. Then $\int_E (h-k) \cdot \mu = 0$. Let $E_1 = \{k < h\}$ and $E_2 = \{k > h\}$. Integrating $h - k$ over these sets it obtains $\mu(E_1) = \mu(E_2) = 0$. Thus $k = h$ μ-almost everywhere. □

The function h obtained in this theorem is called the Radon-Nikodym derivative of ν with respect to μ, and it is usually denoted by $d\nu/d\mu$. The justification for this notation is that it satisfies familiar calculus properties.

Example 7. Let $X = [0, 1]$ and let $f : X \to X$ be a differentiable function whose derivative is bounded and nowhere zero. If μ is Lebesgue measure on X, then $\nu(E) = \mu(f^{-1}E)$ is a measure on X which is absolutely continuous with respect to μ, and the Radon-Nikodym derivative $d\nu/d\mu$ is $|f'(x)|$

6. Remarks and References

The proofs the Lebesgue Decomposition and Radon-Nikodym Theorems given here, using the Riesz Representation Theorem as the main tool, are due to J. von Neumann, see Dudley, *Real Analysis and Probability*, Chapman & Hall, New York, 1989.

There are stronger versions of these theorems. For instance, in the Lebesgue Decomposition it suffices to assume that the measures are σ-finite; in the Radon-Nikodym Theorem it suffices that μ be σ-finite and ν be finite. See Dudley *loc. cit.*