Notes on product measures for Math 501, Fall 2010

Eric A. Carlen[1]
Rutgers University

November 9, 2010

1 Product measures

1.1 Construction of product measure

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \mu)$ be two measure spaces. A *rectangle* in $X \times Y$ is a subset of $X \times Y$ of the form $A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

1.1 LEMMA (The rectangle algebra)**.** *For any two measure spaces* $(X, \mathcal{M}, \mu)$ *and* $(Y, \mathcal{N}, \mu)$*, let* $\mathcal{A}$ *denote the set of all finite disjoint unions of rectangles in* $X \times Y$*. Then* $\mathcal{A}$ *is an algebra, called the* rectangle algebra *in* $X \times Y$*.*

Proof: By the basic lemma on algebras, it suffices to show that the set of rectangles is closed under intersection, and that the complement of any rectangle belongs to $\mathcal{A}$.

For the first point, let $A \times B$ and $C \times D$ be two rectangles. Then

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap C) \ ,$$

and or the second point,

$$(A \times B)^c = (A^c \times B) \cup (X \times B^c) \ .$$

Thus, $\mathring{A}$ is an algebra. □

We next define a premeasure m on $\mathcal{A}$ by setting

$$m\left(\cup_{j=1}^n A_j \times B_j\right) = \sum_{j=1}^n \mu(A_j)\nu(B_j)$$

for any finite disjoint union $\cup_{j=1}^n A_j \times B_j$ of rectangles. It is easy to see, by considering a common refinement of any two representations of a set in $\mathcal{A}$ as a finite disjoint union of rectangles that m does not depend on the representation, and is well-defined premeasure on $\mathcal{A}$.

1.2 DEFINITION (Product sigma-algebra)**.** For any two measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \mu)$, the *product sigma algebra* in $X \times Y$, denoted by $\mathcal{M} \otimes \mathcal{N}$, is the smallest sigma algebra containing the rectangle algebra $\mathcal{A}$.

We shall now use Caratheodory's Theorem to construct a countably additive measure $\mu \otimes \nu$ on $\mathcal{M} \otimes \mathcal{N}$ that agrees with the premeasure m on $\mathcal{A}$.

The key to this is to show that the premeasure m is continuous at the empty set. The following notation will be useful for this purpose and others: For any set $E \subset X \times Y$, and any $x \in X$, the *section of E at x* is the set $E_x \subset Y$ defined by

$$E_x := \{\ y \in Y\ :\ (x,y) \in E\ \}\ .$$

Likewise, the *section of E at y* is the set $E_y \subset X$ defined by

$$E_y := \{\ x \in X\ :\ (x,y) \in E\ \}\ .$$

We note that $(E^c)_x = (E_x)^c$, and that for arbitrary unions and intersections,

$$(\cup_\alpha E_\alpha)_x = \cup_\alpha (E_\alpha)_x \qquad \text{and} \qquad (\cap_\alpha E_\alpha)_x = \cup_\alpha (E_\alpha)_x\ . \tag{1.1}$$

Of course, the same identities hold for y-sections.

Finally, we note that if $A \in \mathcal{A}$, then for each x and y, $E_x \in \mathcal{A}$ and $E_y \in \mathcal{A}$.

1.3 LEMMA (Continuity at the empty set)**.** *The premeasure m on the rectangle algebra $\mathcal{A}$ is continuous at the empty set.*

Proof: Let $\{A_j\}_{j\ge 1}$ be a decreasing sequence of sets in $\mathcal{A}$ with $\cap_{j\ge 1} A_j = \emptyset$ and $m(A_1) < \infty$. We must show that $\lim_{j\to\infty} m(A_j) = 0$.

Let us write

$$A_1 = \cup_{k=1}^n B_k \times C_k$$

and for $j > 1$,

$$A_j = \cup_{i=1}^{n_j} E_{j,i} \times F_{j,i} \tag{1.2}$$

where all unions are disjoint unions of rectangles.

Clearly, it suffices to show that, for each k,

$$\lim_{j\to\infty} m((B_k \times C_k) \cap A_j) = 0\ . \tag{1.3}$$

For each $y \in Y$, $j \ge 1$, and $1 \le i \le n_j$ define

$$f_{j,i}(y) := \mu([(B_k \times C_k) \cap (E_{j,i} \times F_{j,i})]_y) = \begin{cases} \mu(B_k \cap E_{j,i}) & y \in C_k \times F_{j,i} \\ 0 & y \notin\in C_k \times F_{j,i} \end{cases}\ .$$

It follows that $f_{j,i}$ is integrable and that

$$\int_Y f_{j,i} \mathrm{d}\nu = \mu(B_k \cap E_{j,i})\nu(C_k \cap F_{j,i}) = m(B_k \times C_k) \cap (E_{j,i} \times F_{j,i}))\ . \tag{1.4}$$

Next, define $f_j := \sum_{i=1}^{n_j} f_{j,i}$. It follows from (1.1) the disjointness of the union in (1.2) that

$$0 \le f_j(y) = \mu([(B_k \times C_k) \cap A_j]_y) \le \mu(B_k)\ . \tag{1.5}$$

In particular,

$$0 \le f_j \le \mu(B_k) 1_{C_k} \ , \tag{1.6}$$

and the latter function is integrable since both $\mu(B_k)$ and $\nu(C_k)$ are finite.

Again from the fact that the union in (1.2) is disjoint, it follows from (1.4) that

$$\int_Y f_j \mathrm{d}\nu = m((B_k \times C_k) \cap A_j) \ . \tag{1.7}$$

Therefore, proving (1.3) amounts to proving that $\lim_{j\to\infty} \int_Y f_j \mathrm{d}\nu = 0$. By the Dominated Convergence Theorem and (1.6), it then suffices to show that for each y, $\lim_{j\to\infty} f_j(y) = 0$.

However, since $\cap_{j=1}^{\infty} A_j = \emptyset$, it follows that for each $y \in Y$,

$$\cap_{j=1}^{\infty}[(B_k \times C_k) \cap A_j]_y = \emptyset \ ,$$

and then since μ is a countably additive measure, and since $\mu(B_k) < \infty$,

$$\lim_{j\to\infty} \mu([(B_k \times C_k) \cap A_j]_y) = 0 \ .$$

By (1.5) this shows that $\lim_{j\to\infty} f_j(y) = 0$. □

Now let us make the further assumption that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are sigma finite. Then, clearly there is a sequence of rectangles $\{A_n\}$ with $m(A_n) < \infty$ for all n and $\cup_{n=1}^{\infty} A_n = X \times Y$. It is then clear that if E is any set in $\mathcal{A}$ with $m(E) = \infty$, and r any positive number, there is an n such that $m(E \cap A_n) > r$.

Therefore, our basic scheme for extending a premeasure on an algebra to countably additive measures on the on the smallest sigma algebra containing the algebra yields us a countably additive measure $\mu \otimes \nu$ on the product sigma algebra $\mathcal{M} \otimes \mathcal{N}$ that extends m on $\mathcal{A}$. This is the *product measure on $X \times Y$ generated by μ and ν.*

1.2 Minkowski's inequality for L^2

Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f_1, \dots, f_n \in L^2(X, \mathcal{M}, \mu)$. Then by the triangle inequality,

$$\left\| \sum_{j=1}^{n} f_j \right\|_2 \le \sum_{j=1}^{n} \|f_j\|_2 \ .$$

Integrals are a generalization of sums, and so it may come as no surprise that there is a generalization of this in which the sum is replaced by an integral.

1.4 THEOREM (Minkowski's inequality for L^2)**.** *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be sigma finite measure space. Let f be an $\mathcal{M} \otimes \mathcal{N}$ measurable function on $X \times Y$ such that*

$$\int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) < \infty \ .$$

Then for almost every y, $f(\cdot, y)$ is in $L^1(X, \mathcal{M}, \mu)$, and

$$\left(\int_Y \left(\int_X f(x,y) \mathrm{d}\mu(x) \right)^2 \mathrm{d}\nu(y) \right)^{1/2} \le \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ . \tag{1.8}$$

Moreover, there is equality in (1.8) if and only if there are function $c \in L^1(X, \mathcal{M}, \mu)$ and $g \in L^2(Y, \mathcal{N}, \nu)$ such that $f(x,y) = c(x)g(y)$ for $\mu \otimes \nu$ almost every (x,y).

Proof: *Step 1: We show that for any $h \in L^2(Y, \mathcal{N}, \nu)$, $h(y)f(x,y) \in L^1(X \times x, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$.* Let $h \in L^2(Y, \mathcal{N}, \nu)$. Then by the Cauchy-Schwarz inequality,

$$\int_X \left(\int_Y |h(y)||f(x,y)| \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) \le \|h\|_2 \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ .$$

By the Fubini-Tonelli Theorem, it follows that $h(y)f(x,y) \in L^1(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$, and that

$$\begin{aligned} \int_Y h(y) \left(\int_X f(x,y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y) &= \int_X \left(\int_Y h(y) f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) \\ &\le \|h\|_2 \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ . \end{aligned} \tag{1.9}$$

Step 2: We approximate $g(y) := \int_X f(x,y)\mathrm{d}\mu(x)$ by square integrable functions.

Let $g(y) := \int_X f(x,y)\mathrm{d}\mu(x)$. Since $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are sigma finite measure spaces, there exists a sequence of rectangles $\{A_n\}$ with $m(A_n) < \infty$ for all n and $\cup_{n=1}^{\infty} A_n = X \times Y$.

Define $E_n := \{ \ x \ : \ |g(x)| \le n \ \} \cap A_n$ and $g_n = 1_{E_n} g^*$, so that $\|g_n\|_2 < \infty$. Note that fro all y, $|g_n(y)|^2 = g_n(y)g(y)$. Then with $h := g_n$, (1.9) says that

$$\|g_n\|_2^2 = \int_Y g_n g \mathrm{d}\nu \le \|g_n\|_2 \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ .$$

Since $\|g_n\|_2$ is finite, if $\|g_n\|_2 \neq 0$, we may divide through by $\|g_n\|_2$ to obtain

$$\|g_n\|_2 \le \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ . \tag{1.10}$$

If $\|g_n\|_2 = 0$, (1.10) is trivially true.

By the Monotone Convergence Theorem, $\lim_{n\to\infty} \|g_n\|_2 = \|g\|_2$, and hence it follows from (1.10) that

$$\|g\|_2 \le \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \ . \tag{1.11}$$

This is equivalent to the inequality (1.8).

Step 3: Now that the inequality itself is proved, we address the cases of equality: We now know that $\|g\|_2 < \infty$, and may repeat the argument of Step 2, but this time without the approximation:

Suppose that equality holds in (1.11) and that the right side is finite and non-zero. Then by the definition of g, and the Fubini-Tonelli Theorem and the Cauchy-Schwarz inequality,

$$\begin{aligned} \|g\|_2^2 &= \int_Y g^*(y) \left(\int_X f(x,y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y) \\ &= \int_X \left(\int_Y g^*(y) f(x,y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) \\ &\le \|g\|_2 \int_X \left(\int_Y |f(x,y)|^2 \mathrm{d}\nu(y) \right)^{1/2} \mathrm{d}\mu(x) \\ &= \|g\|_2^2 \end{aligned}$$

It follows from the condition for equality in the Cauchy-Schwarz inequality that for μ almost every x, g and $f(x,\cdot)$ are linearly dependent. Then there is a constant $c(x)$ such that $f(x,y) = c(x)g(y)$ for almost every y. Since then $c(x) = \|g\|_2^{-2}\int_Y f(x,y)g(y)\mathrm{d}(y)$, c is measurable, and then it readily follows that c is integrable. □

1.3 Convolution operators on $L^2(\mathbb{R}^n, \mathcal{B}, \mu)$

We now specialize to the case in which $X = \mathbb{R}^n$, $\mathcal{B}$ is its Borel sigma algebra, and μ is Lebesgue measure. Since $\mathbb{R}^n$ is an abelian group, and since Lebesgue measure is invariant under the group action (translation), we have some additional structure that allows us to define an important class of operators on $L^2(\mathbb{R}^n, \mathcal{B}, \mu)$ – convolution operators.

In fact, the arguments that follow are easily seen to apply in other abelian groups as well, and in particular in the case of S^1. Therefore, we will keep our notation somewhat general.

Let $g \in L^1(X, \mathcal{M}, \mu)$, and let $h \in L^2(X, \mathcal{M}, \mu)$. Then define

$$f(x,y) := h(x-y)g(y) \ .$$

Then

$$\int_X \left(\int_X |f(x,y)|^2 \mathrm{d}\mu(x)\right)^{1/2} \mathrm{d}\mu(y) = \|g\|_1 \|h\|_2 \ .$$

Therefore, by Minkowski's Inequality, $h(x-y)g(y)$ is integrable in y for almost every x, and moreover,

$$g * h(x) := \int_X h(x-y)g(y)\mathrm{d}\mu(y)$$

is in $L^2(X, \mathcal{M}, \mu)$, and

$$\|g * h\|_2 \leq \|g\|_1 \|h\|_2 \ . \tag{1.12}$$

Then, since integration is linear, the map $C_g : h \mapsto g * h$ is a linear transformation from $L^2(X, \mathcal{M}, \mu)$ to $L^2(X, \mathcal{M}, \mu)$. The ineuqality shows that C_g is continuous: Indeed,

$$\|C_g(h_1) - C_g(h_2)\|_2 = \|C_g(h_1 - h_2)\|_2 \leq \|g\|_1 \|h_1 - h_2\|_2 \ .$$

For example, consider the case in which X is replaced by S^1 and μ is Lebesgue measure on S^1. For any given N, ket F_N denote the Fejer kernel of degree N. Since $\|F_N\|_1 = 1$, we see that

$$\|F_N * h_1 - F_N * h_2\|_2 \leq \|h_1 - h_2\|_2 \ .$$

The reason that convolution operators are so important is that they are essentially the only continuous linear transformations on $L^2(X, \mathcal{M}, \mu)$ that commute with translations.

1.4 Orthonormal bases in $L^2(\mu \otimes \nu)$

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two sigma-finite measure spaces. Suppose also that both $L^2(X, \mathcal{M}, \mu)$ and $L^2(Y, \mathcal{N}, \mu)$ are separable, so that both posses orthonormal bases. In practice, this means that both $\mathcal{M}$ and $\mathcal{N}$ are the sigma algebras generated by some countable algebra of sets.

Let $\{\varphi_i\}_{i\geq 1}$ be an orthonormal basis for $L^2(X,\mathcal{M},\mu)$, and let $\{\psi_j\}_{j\geq 1}$ be an orthonormal basis for $L^2(Y,\mathcal{N},\mu)$. Let $\varphi_i \otimes \psi_j$ denote the function on $X \times Y$ given by

$$\varphi_i \otimes \psi_j(x,y) = \varphi_i(x)\psi_j(y)\ .$$

Then by the Fubini-Tonelli Theorem,

$$\int_{X\times Y} (\varphi_i \otimes \psi_j)^*(\varphi_k \otimes \psi_\ell)\mathrm{d}\mu\otimes\nu = \langle \varphi_i, \varphi_k\rangle_{L^2(X,\mathcal{M},\mu)}\langle \psi_j, \psi_\ell\rangle_{L^2(Y,\mathcal{N},\mu)} = \begin{cases} 1 & i = k \text{ and } j = \ell \\ 0 & \text{otherwise} \end{cases}\ .$$

Therefore, $\{\varphi_i \otimes \psi_j\}_{i,j\geq 0}$ is orthonormal in $L^2(X\times Y, \mathcal{M}\otimes\mathcal{N}, \mu\otimes\nu)$. Is it total, and hence an orthonormal basis for $L^2(X\times Y, \mathcal{M}\otimes\mathcal{N}, \mu\otimes\nu)$? Yes, always:

1.5 THEOREM (Products of orthonormal bases)**.** *Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\mu)$ be two sigma-finite measure spaces, and suppose that $\{\varphi_i\}_{i\geq 1}$ is an orthonormal basis for $L^2(X,\mathcal{M},\mu)$, and that $\{\psi_j\}_{j\geq 1}$ is an orthonormal basis for $L^2(Y,\mathcal{N},\mu)$. Then $\{\varphi_i \otimes \psi_j\}_{i,j\geq 0}$ is an orthonormal basis for $L^2(X\times Y, \mathcal{M}\otimes\mathcal{N}, \mu\otimes\nu)$.*

We first prove the following lemma:

1.6 LEMMA. *Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\mu)$ satisfy the conditions of Theorem 1.5, and let $A\times B$ be any rectangle with finite product measure. Then for every $\epsilon > 0$, there is an $N > 0$ and numbers $c_{i,j}$, $1 \leq i,j \leq N$ such that with $g := \sum_{i,j\leq N} c_{i,j}\varphi_i\otimes\psi_j$, $\|1_{A\times B} - g\|_2 < \epsilon$.*

Proof: $1_{A\times B}(x,y) = 1_A(x)1_B(y)$. We may suppose both $\mu(A)$ and $\nu(B)$ are strictly positive, or else there is nothing to prove. Let

$$a_n = \int_X \varphi_n^* 1_A \mathrm{d}\mu \qquad \text{and} \qquad b_n = \int_Y \psi_n^* 1_B \mathrm{d}\mu\ .$$

Then since $\{\varphi_n\}$ and $\{\psi_n\}$ are orthonormal bases, there is an $N < \infty$ so that

$$\|1_A - \sum_{n=1}^N a_n\varphi_n\|_2 < \frac{\epsilon}{2(\nu(B))^{1/2}} \qquad \text{and} \qquad \|1_B - \sum_{n=1}^N b_n\psi_n\|_2 < \frac{\epsilon}{2(\mu(A))^{1/2}}\ .$$

Define

$$g(x,y) = \left(\sum_{n=1}^N a_n\varphi_n(x)\right)\left(\sum_{n=1}^N b_n\psi_n(y)\right)\ .$$

Then

$$|1_{A\times B}(x,y) - g(x,y)| = \left|1_A(x) - \sum_{n=1}^N a_n\varphi_n(x)\right| 1_B(y) + \left|\sum_{n=1}^N a_n\varphi_n(x)\right|\left|1_B(x) - \sum_{n=1}^N b_n\psi_n(x)\right|\ ,$$

and hence, by Bessels's inequality, which tells us that $\|\sum_{n=1}^N a_n\varphi_n\|_2 \leq \|1_A\|_2 = (\mu(A))^{1/2}$,

$$\|1_{A\times B} - g\|_2 \leq \|1_A - \sum_{n=1}^N a_n\varphi_n\|_2(\nu(B))^{1/2} + \|1_B - \sum_{n=1}^N b_n\psi_n\|_2(\mu(A))^{1/2} < \epsilon\ .$$

Proof of Theorem 1.5: We must show that $\{\varphi_i \otimes \psi_j\}_{i,j\geq 0}$ is total. Therefore, suppose that $f \in L^2(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$ is orthogonal to each $\varphi_i \otimes \psi_j$.

Pick $\epsilon > 0$. For $n \in \mathbb{N}$, define

$$E_n := \{(x,y) \;\; 1/n \leq f(x,y) \leq n \;\} \ ,$$

and $f_n := 1_{E_n} f$. By the Dominated Convergence Theorem,

$$\lim_{n\to\infty} \|f_n - f\|_2 = 0 \ .$$

Hence for some finite n, $\|f - f_n\|_2 < \epsilon$.

Note that $(1/n)|f_n| \leq |f|^2$ and hence

$$\int_{X\times Y} |f_n| \mathrm{d}(\mu \otimes \nu) \leq n\|f\|_2^2 < \infty \ .$$

Thus, $f_n \in L^1(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$.

By the theorem on approximation by really simple functions, it follows that that is a really simple function h such that $\|f_n - h\|_1 \leq \epsilon^2/(2n)$, and such that ess sup$|h|$ = ess sup$|f_n|$, and in the case at hand ess sup$|f_n| = n$. Therefore,

$$\|f_n - h\|_2^2 \leq 2n \int_{X\times Y} |f_n - h| \mathrm{d}\mu \otimes \nu = 2n\|f_n - h\|_1 \ ,$$

and hence $\|f_n - h\|_2 \leq \epsilon$.

Thus, $\|f - h\|_2 \leq 2\epsilon$. However, h is a finite linear combination of characteristic functions of rectangles. Therefore, by (1.6) and the triangle inequality, there is an $N > 0$ and numbers $c_{i,j}$, $1 \leq i, j \leq N$ such that with $g := \sum_{i,j\leq N} c_{i,j}\varphi_i \otimes \psi_j$, $\|h - g\|_2 < \epsilon$. Finally, we have

$$\|f - g\|_2 \leq 3\epsilon \ .$$

By hypothesis, $\langle f, g\rangle = 0$, and hence

$$\begin{aligned} \|f\|_2^2 &= \langle f, f\rangle = \langle f, f - g\rangle + \langle f, g\rangle \\ &\leq \|f\|_2\|f - g\|_2 \leq \|f\|_2 3\epsilon \ . \end{aligned}$$

It follows that $\|f\|_2 \leq 3\epsilon$. Since $\epsilon > 0$ is arbitrary, $\|f\|_2 = 0$, and hence $f = 0$. □

1.5 Hilbert-Schmidt operators from $L^2(X, \mathcal{M}, \mu)$ to $L^2(Y, \mathcal{N}, \nu)$

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be sigma-finite measure spaces. Let $K(x,y) \in L^2(X\times Y, \mathcal{M}\otimes\mathcal{N}, \mu\otimes\nu)$. Then for almost evert x, $K(x,\cdot) \in L^2(Y, \mathcal{N}, \nu)$, and so for almost every x, and every $f \in L^2(Y, \mathcal{N}, \nu)$, $K(x,\cdot)f(\cdot) \in L^1(Y, \mathcal{N}, \nu)$. We may therefore define the function

$$Kf(x) := \int_Y K(x,y)f(y)\mathrm{d}\nu(y) \ .$$

We claim that for any $g \in L^2(X, \mathcal{M}, \mu)$, $gKf \in L^1(X, \mathcal{M}, \mu)$. To see that note that

$$|Kf(x)| \leq \int_X |K(x,y)||f(y)|\mathrm{d}\mu(y) \ ,$$

and so by the Fubini-Tonelli Theorem and the Cauchy-Schwarz Inequality

$$\begin{aligned}\int_X |g(x)||Kf(x)|\mathrm{d}\mu(x) &\leq \int_X |g(x)| \left(\int_X |K(x,y)||f(y)|\mathrm{d}\mu(y)\right)\mathrm{d}\mu(x)\\ &= \int_{X\times X} |g(x)||f(y)||K(x,y)|\mathrm{d}\mu\otimes\nu(x,y)\\ &\leq \|g\otimes f\|_2\|K\|_2 = \|g\|_2\|f\|_2\|K\|_2 \ .\end{aligned}$$

Now that we know gKf is integrable, we apply the Fubini-Tonelli Theorem once more to conclude

$$\begin{aligned}\int_X g(x)Kf(x)\mathrm{d}\mu(x) &= \int_X g(x) \left(\int_X K(x,y)f(y)\mathrm{d}\mu(y)\right)\mathrm{d}\mu(x)\\ &= \int_{X\times X} g(x)f(y)K(x,y)\mathrm{d}\mu\otimes\nu(x,y)\\ &\leq \|g\otimes f\|_2\|K\|_2 = \|g\|_2\|f\|_2\|K\|_2 \ .\end{aligned}$$

Now we make one of our standard arguments: Let $\{A_n\}$ be s sequence of measurable subsets of X with $\mu(A_n) < \infty$ for all n, and $\cup_{n=1}^{\infty} A_n = X$. Define $E_n := \{\ x\ :\ |Kf(x)| \leq n\ \} \cap A_n$, and let $g_n = 1_{E_n}(Kf)^*$. Then for each n, $g_n \in L^2(X,\mathcal{M},\mu)$ and by the Monotone Convergence Theorem, $\lim_{n\to\infty}\|g_n\|_2 = \|Kf\|_2$.

Thus,

$$\|g_n\|_2^2 = \int\int_X g_n(x)Kf(x)\mathrm{d}\mu(x) \leq \|g_n\|_2\|f\|_2\|K\|_2 \ .$$

We conclude that $\|g_n\|_2 \leq \|K\|_2\|f\|_2$. Taking the limit $n\to\infty$, we see that

$$\|Kf\|_2 \leq \|K\|_2\|f\|_2 \ . \tag{1.13}$$

Thus, the map $K : f \mapsto Kf$ is a transformation from $L^2(Y,\mathcal{N},\nu)$ into $L^2(X,\mathcal{M},\mu)$. Since integration is linear, it is a linear transformation from $L^2(Y,\mathcal{N},\nu)$ into $L^2(X,\mathcal{M},\mu)$. Then, the inequality (1.13) implies that K is continuous. Indeed,

$$\|Kf_1 - Kf_2\|_2 = \|K(f_1-f_2)\|_2 \leq \|K\|_2\|f_1-f_2\|_2 \ .$$

It turns out that K has a much stronger property than continuity: The image of the unit ball in $L^2(Y,\mathcal{N},\nu)$ under K is contained in a compact subset of $L^2(X,\mathcal{M},\mu)$. This brings us to an important result:

1.7 THEOREM (Compactness property of Hilbert-Schmidt operators)**.** *Let K be a Hilbert-Schmidt operator from $L^2(Y,\mathcal{N},\nu)$ into $L^2(X,\mathcal{M},\mu)$. Then the closure of*

$$\{\ Kf\ :\ \|f\|_2 \leq 1\ \}$$

is a compact subset of $L^2(X,\mathcal{M},\mu)$.

Before proving this, we prove some simple lemmas on compactness that are of general utility.

1.8 DEFINITION (Totally bounded sets)**.** Let (X,d) be a metric space. An *ϵ-cover* of a set $A \subset X$ is a collection of subsets of X having diameter at most ϵ whose union includes A. A set $A \subset X$ is *totally bounded* if for every $\epsilon > 0$, there is a finite ϵ-cover of A.

1.9 THEOREM (Compactness, completeness and total boundedness). *Let (X,d) be a metric space. Then X is compact if and only if X is complete and totally bounded.*

Proof: Suppose X is compact. The completeness of X follows easily from the equivalence of compactness and sequential compactness for metric spaces: Every Cauchy sequence has a convergent subsequence, but a Cauchy sequence with a convergence subsequence is itself convergent. Hence X is complete. On the other hand, for each $\epsilon > 0$, $\{B_{\epsilon/2}(x) \ : \ x \in X\}$ is an open ϵ cover of X. Since X, is compact, there exists a finite subcover. Thus, X is totally bounded.

Now suppose that X is complete and totally bounded. Let $\{x_n\}$ be an arbitrary sequence in X. By the equivalence of compactness and sequential compactness for metric spaces, it suffices to show that $\{x_n\}$ has a convergent subsequence. Consider any finite 1-cover of X. At least one of the sets in the cover contains infinitely many terms in the sequence $\{x_n\}$. Chose one such set, and define a new sequence $\{x_n^{(1)}\}$ by discarding all terms that are not in the chosen set. The new sequence is a subsequences of the original sequence.

Now proceed inductively: For each natural number k, suppose $\{x_n^{(k)}\}$ is already defined. Consider any finite $1/(k+1)$-cover of X, and choose one set in the cover that contains infinitely many terms from the sequence $\{x_n^{(k)}\}$. Define $\{x_n^{(k+1)}\}$ by discarding all terms that are not in the chosen set. Note that $\{x_n^{(k+1)}\}$ is a subsequence of $\{x_n^{(k)}\}$, and hence, by induction, of the original sequence.

Finally, consider the Cantor diagonal sequence $\{x_n^{(k)}\}$. By construction, this is a Cauchy subsequence of the original sequence. By the completeness of X, it converges to some $x \in X$. Thus, every sequence $\{x_n\}$ has a convergent subsequence. □

1.10 LEMMA (Compactness Criterion). *Let (X,d) be a complete metric space. Suppose that for each $\epsilon > 0$, there is a metric space (W, d_W), a mapping $\Phi : X \to W$, and a $\delta > 0$ so that $\Phi(X)$ is totally bounded and whenever $x, y \in X$ are such that $d_W(\Phi(x), \Phi(y)) < \delta$, then $d(x,y) < \epsilon$. Then X is compact.*

Proof: We must show that X is totally bounded. Pick $\epsilon > 0$. Let (W, d_W), δ and Φ be as in Lemma 1.10. Let $\{V_1, \dots, V_n\}$ be any δ-cover of $\Phi(X)$. Then $\{\Phi^{-1}(V_1), \dots, \Phi^{-1}(V_n)\}$ is an *epsilon*-cover of X. □.

Note that in Lemma 1.10, the metrics space (W, d_W) and the map Φ may change with ϵ. This will be the case in our first application.

Proof of Theorem 1.7: For any positive integer N, define a map $\Phi_N : L^2(X, \mathcal{M}, \mu) \to \mathbb{C}^N$ by

$$\Phi_N(h) = (\langle \varphi_1, h\rangle, \dots, \langle \varphi_N, h\rangle) \ .$$

Note that with $\|\cdot\|$ denoteing the norm on $\mathbb{C}^N$, Bessel's inequality gives us

$$\|\Phi_N(h)\| = \left(\sum_{j=1}^N |\langle \varphi_j, h\rangle|^2\right)^{1/2} \le \|h\|_2 \ . \tag{1.14}$$

Define $Z := \{Kf \ : \ \|f\|_2 \le 1 \ \}$, and equip Z with the metric it inherits as a subset of $L^2(X, \mathcal{M}, \mu)$ By (1.13), if f is in the unit ball in $L^2(Y, \mathcal{N}, \nu)$, $\|Kf\|_2 \le \|K\|_2$. Together with (**??**) this gives us

$$\|\Phi_N(Kf)\|_2 \le \|K\|_2 \ .$$

In particular, $\Phi_N(Z)$ is a bounded, and hence totally bounded subset of $\mathbb{C}^N$, which is a complete metric space.

Now et $\epsilon > 0$ be given. We will show for an appropriate choice of N, there is a $\delta > 0$ so that if f and g belong to the unit ball in $L^2(Y, \mathcal{N}, \nu)$ and $\|\Phi_N(Kf) - \Phi_N(Kg)\|_2 \leq \delta$, then $\|f - g\|_2 < \epsilon$. Since $\mathbb{C}^N$ is complete, it will then follow from Lemma 1.10 that the closure of Z is compact.

We first make a finite-rank approximation of the linear transformation K: Since $\int_{X \times Y} |K(x,y)|^2 \mathrm{d}\mu \otimes \nu(x,y) < \infty$, if we define

$$K_N(x,y) = \sum_{m,n \leq N} \left(\int_{X \times Y} K(x,y) \varphi_m(x) \psi_n(y) \mathrm{d}\mu \otimes \nu(x,y) \right) \varphi_m \otimes \psi_n(x,y) \ ,$$

the fact that $\{\varphi_m \otimes \psi_n\}_{m,n}$ is an orthonormal basis tells us that

$$\lim_{N \to \infty} \|K - K_N\|_{L^2(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)} = 0 \ . \tag{1.15}$$

The linear transformation sending f to $\int_Y K_N(x,y) f(y) \mathrm{d}\nu(y)$ is not only Hilbert-Schmidt; it is finite rank: $\{\varphi_1, \ldots, \varphi_N\}$ is an orthonormal basis for the range. It follows from (1.13) and (1.15) that for some finite N,

$$\|K_N f - Kf\|_2 \leq \frac{\epsilon}{6} \|f\|_2 \ . \tag{1.16}$$

Now note that for any $f \in L^2(Y, \mathcal{N}, \nu)$,

$$K_N f = \sum_{j=1}^{N} (\Phi_N(K_N f))_j \, \varphi_j \qquad \text{and thus} \qquad \|K_N f\|_2 = \|\Phi_N(K_N f)\| \ , \tag{1.17}$$

where the norm on the right is the norm in $\mathbb{C}^N$.

Now consider any f and g in the unit ball in $L^2(Y, \mathcal{N}, \nu)$. By the linearity of K, K_N and Φ_N, and (1.16) we have

$$\begin{aligned} \|Kf - Kg\|_2 = \|K(f-g)\|_2 &\leq \|K(f-g) - K_N(f-g)\|_2 + \|K_N(f-g)\|_2 \\ &\leq \frac{\epsilon}{6} \|f - g\|_2 + \|K_N(f-g)\|_2 \end{aligned} \tag{1.18}$$

Next, by (1.17), and then (1.16)

$$\begin{aligned} \|K_N(f-g)\|_2 = \|\Phi_N[K_N(f-g)]\| &\leq \|\Phi_N[K(f-g) - K_N(f-g)]\| + \|\Phi_N[K(f-g)]\| \\ &= \|K(f-g) - K_N(f-g)\|_2 + \|\Phi_N[K(f-g)]\| \\ &\leq \frac{\epsilon}{6} \|f - g\|_2 + \|\Phi_N[K(f-g)]\| \ . \end{aligned}$$

Altogether, since $\|f - g\|_2 \leq 2$,

$$\|K(f-g)\|_2 \leq \frac{2}{3}\epsilon + \|\Phi_N[K(f-g)]\| \ .$$

Thus,

$$\|\Phi_N[Kf] - \Phi_N[Kg]\| \leq \frac{1}{3}\epsilon \quad \Rightarrow \quad \|Kf - Kg\|_2 < \epsilon \ .$$

We see that choosing N so that (1.16) is satisfied, and then taking $\delta = \epsilon/3$, the conditions of Lemma 1.10 are met. □