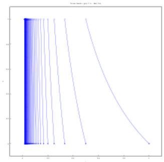
Absolutely Continuous Measures

A transformation may have many invariant measures, but some are more improtant than others.

Example 1: To find the digits in the standard continued fraction expansion of a real number x:

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}} = [a_o : a_1, a_2, a_3, \dots]$$

we need the Gauß map $G : [0,1) \to (0,1]$:



$$G(x) = \frac{1}{x} - \lfloor \frac{1}{x} \rfloor.$$

Absolute Continuous Measures

$$a_0 = \lfloor x \rfloor \qquad x_1 = x - a_0.$$
$$a_i = \lfloor 1/x_i \rfloor \qquad x_{i+1} = G(x_i) \quad \text{(stop if } x_{i+1} = 0\text{)}.$$

For example:

$$\pi = [3; 7, 15, 1, 292, 1, 1, 1, \dots]$$

and

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, \dots]$$

Lebesgue measure is not G-invariant.

Absolutely Continuous Measures

Definition: A measure μ is called absolutely continuous w.r.t. the measure ν (notation: $\mu \ll \nu$) if $\nu(A) = 0$ implies $\mu(A) = 0$. If both $\mu \ll \nu$ and $\nu \ll \mu$, then μ and ν are called equivalent.

Theorem of Radon-Nikodym: If μ is a probability measure and $\mu \ll \nu$ then there is a function $h \in L^1(\nu)$ (called Radon-Nikodym derivative or density) such that $\mu(A) = \int_A h(x)\, d\nu(x)$ for every measurable set A.

Notation: $h(x) = \frac{d\mu(x)}{d\nu(x)}$.

Example 1: Lebesgue measure is not G-invariant, but there is a probability measure μ_G that is absolutely continuous w.r.t. Lebesgue.

$$\mu_G(A) = \int_A h(x)\, dx \quad \text{for} \quad h(x) = \frac{1}{\log 2} \frac{1}{1+x}.$$

Absolutely Continuous Measures

Proposition 1. Suppose that $\mu \ll \nu$ are both T-invariant probability measures, with a common σ-algebra $\mathcal{B}$ of measurable sets. If ν is ergodic, then $\mu = \nu$.

Proof: First we show that μ is ergodic. Indeed, otherwise there is a T-invariant set A such that $\mu(A) > 0$ and $\mu(A^c) > 0$. By ergodicity of ν at least one of A or A^c must have ν-measure 0, but this would contradict that $\mu \ll \nu$.

Now let $A \in \mathcal{B}$ and let $Y \subset X$ be the set of ν-typical points. Then $\nu(Y^c) = 0$ and hence $\mu(Y^c) = 0$. Applying Birkhoff's Ergodic Theorem to μ and ν separately for $\psi = 1_A$ and some μ-typical $y \in Y$, we get

$$\mu(A) = \lim_n \frac{1}{n} \sum_{i=0}^{n-1} \psi \circ T(y) = \nu(A).$$

But $A \in \mathcal{B}$ was arbitrary, so $\mu = \nu$.

Absolutely Continuous Measures

Exercise 4.1: Show that the condition of ergodicity is essential for Proposition 1?

If $\mu \ll \nu$ and μ is ergodic. Does it follow that $\mu = \nu$?

Proposition 2: Let $T : U \subset \mathbb{R}^n \to U$ be (piecewise) differentiable, and μ is absolutely continuous w.r.t. Lebesgue. Then μ is T-invariant if and only if its density $h = \frac{d\mu}{dx}$ satisfies

$$h(x) = \sum_{T(y)=x} \frac{h(y)}{|\det DT(y)|} \tag{1}$$

for every x.

Absolutely Continuous Measures

Proof of Proposition 2: The T-invariance means that $d\mu(x) = d\mu(T^{-1}(x))$, but we need to be aware that T^{-1} is multivalued. So it is more careful to split the space U into pieces U_n such that the restrictions $T_n := T|U_n$ are diffeomorphic (onto their images) and write $y_n = T_n^{-1}(x) = T^{-1}(x) \cap U_n$. Then we obtain (using the change of coordinates)

$$\begin{aligned} h(x)\ dx &= d\mu(x) = d\mu(T^{-1}(x)) = \sum_n d\mu \circ T_n^{-1}(x) \\ &= \sum_n h(y_n) |\det(DT_n^{-1})(x)| dy_n = \sum_n \frac{h(y_n)}{\det |DT(y_n)|} dy_n. \end{aligned}$$

Conversely, if (1) holds, then the above computation gives $d\mu(x) = d\mu \circ T^{-1}(x)$.

Absolutely Continuous Measures

Example 1 continued: The Gauß map has invariant density $h(x) = \frac{1}{\log 2}\frac{1}{1+x}$. Here $\frac{1}{\log 2}$ is just the normalising factor (so that $\int_0^1 h(x)dx = 1$).

Let $I_n = (\frac{1}{n+1}, \frac{1}{n}]$ for $n = 1, 2, 3, \dots$ be the domains of the branches of G, and for $x \in (0,1)$, and $y_n := G^{-1}(x) \cap I_n = \frac{1}{x+n}$. Also $G'(y_n) = -\frac{1}{y_n^2}$. Therefore

$$\begin{aligned}
\sum_{n\geq 1} \frac{h(y_n)}{|G'(y_n)|} &= \frac{1}{\log 2}\sum_{n\geq 1}\frac{y_n^2}{1+y_n} = \frac{1}{\log 2}\sum_{n\geq 1}\frac{\frac{1}{(x+n)^2}}{1+\frac{1}{x+n}} \\
&= \frac{1}{\log 2}\sum_{n\geq 1}\frac{1}{x+n}\cdot\frac{1}{x+n+1} \\
&= \frac{1}{\log 2}\sum_{n\geq 1}\frac{1}{x+n}-\frac{1}{x+n+1} \qquad \text{telescoping series} \\
&= \frac{1}{\log 2}\frac{1}{x+1} = h(x).
\end{aligned}$$

Absolutely Continuous Measures

Exercise 4.2: Compute the average frequency of the digit 1 for points that are normal w.r.t. the standard continued fraction.

Exercise 4.3: Show that for each integer $n \geq 2$, the interval map given by

$$T_n(x) = \begin{cases} nx & \text{if } 0 \leq x \leq \frac{1}{n}, \\ \frac{1}{x} - \lfloor \frac{1}{x} \rfloor & \text{if } \frac{1}{n} < x \leq 1, \end{cases}$$

has invariant density $h(x) = \frac{1}{\log 2} \frac{1}{1+x}$.

Absolutely Continuous Measures

Example 2: The map $T : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $T(x) = x - \frac{1}{x}$ is called the Boole transformation. It is 2-to-1; the two preimages of $x \in \mathbb{R}$ are $y_{\pm} = \frac{1}{2}(x \pm \sqrt{x^2+4})$. Clearly $T'(x) = 1 + \frac{1}{x^2}$. It can be shown that

$$\frac{1}{|T'(y_-)|} + \frac{1}{|T'(y_+)|} = 1.$$

Indeed,

$$|T'(y_{\pm})| = 1 + \frac{2}{x^2 + 2 \pm x\sqrt{x^2+4}}$$

and

$$\frac{1}{|T'(y_{\pm})|} = \frac{x^2 + 2 \pm x\sqrt{x^2+4}}{x^2 + 4 \pm x\sqrt{x^2+4}},$$

and

Absolutely Continuous Measures

$$
\begin{aligned}
& \frac{1}{|T'(y_-)|} + \frac{1}{|T'(y_+)|} \\
= \quad & \frac{x^2+2-x\sqrt{x^2+4}}{x^2+4-x\sqrt{x^2+4}} + \frac{x^2+2+x\sqrt{x^2+4}}{x^2+4+x\sqrt{x^2+4}} \\
= \quad & \frac{(x^2+2-x\sqrt{x^2+4})(x^2+4+x\sqrt{x^2+4})}{(x^2+4)^2-x^2(x^2+4)} \\
& \quad + \frac{(x^2+2+x\sqrt{x^2+4})(x^2+4-x\sqrt{x^2+4})}{(x^2+4)^2-x^2(x^2+4)} \\
= \quad & \frac{(x^2+2)^2-x^2(x^2+4)+2(x^2+2)-2x\sqrt{x^2+4}}{4(x^2+4)} + \\
& \quad \frac{(x^2+2)^2-x^2(x^2+4)+2(x^2+2)+2x\sqrt{x^2+4}}{4(x^2+4)} \\
= \quad & \frac{4(x^2+2)+8}{4(x^2+4)} = 1.
\end{aligned}
$$

Therefore $h(x) \equiv 1$ is an invariant density, so Lebesgue measure is

The Folklore Theorem

Example 3: If $T : [0,1] \to [0,1]$ is (countably) piecewise linear, and each branch $T : I_n \to [0,1]$ (on which T is affine) is onto, then T preserves Lebesgue measure. Indeed, the intervals I_n have pairwise disjoint interiors, and their lengths add up to 1. If s_n is the slope of $T : I_n \to [0,1]$, then $s_n = 1/|I_n|$. Therefore

$$\sum_n \frac{1}{DT(y_n)} = \sum_n \frac{1}{s_n} = \sum_n |I_n| = 1.$$

"Folklore" Theorem If $T : \mathbb{S}^1 \to \mathbb{S}^1$ is a C^2 expanding circle map, then it preserves a measure μ equivalent to Lebesgue, and μ is ergodic.

Expanding means that there is $\lambda > 1$ such that $|T'(x)| \geq \lambda$ for all $x \in \mathbb{S}^1$. The above theorem can be proved in more generality, but in the stated version it conveys the ideas more clearly.

The Folklore Theorem

Proof: Using the Mean Value Theorem twice, we obtain

$$\begin{aligned}\log \frac{|T'(x)|}{|T'(y)|} &= \log(1 + \frac{|T'(x)| - |T'(y)|}{|T'(y)|}) \leq \frac{|T'(x)| - |T'(y)|}{|T'(y)|} \\ &= \frac{|T''(\xi)| \cdot |x-y|}{|T'(y)|} = \frac{|T''(\xi)|}{|T'(y)|} \frac{|Tx - Ty|}{T'(\zeta)}.\end{aligned}$$

Since T is expanding, the denominators are $\geq \lambda$ and since T is C^2 on a compact space, also $|T''(\xi)|$ is bounded. Therefore there is some $K \leq \sup |T''(\xi)|/\lambda^2$ such that

$$\log \frac{|T'(x)|}{|T'(y)|} \leq K|T(x) - T(y)|.$$

The chain rule then gives:

$$\log \frac{|DT^n(x)|}{|DT^n(y)|} = \sum_{i=0}^{n-1} \log \frac{|T'(T^i x)|}{|T'(T^i y)|} \leq K \sum_{i=1}^{n} |T^i(x) - T^i(y)|.$$

The Folklore Theorem

Since T is a continuous expanding map of the circle, it wraps the circle d times around itself, and for each n, there are d^n pairwise disjoint intervals Z_n such that $T^n : Z_n \to \mathbb{S}^1$ is onto, with slope at least λ^n. If we take x, y above in one such Z_n, then

$$|x - y| \leq \lambda^{-n} |T^n(x) - T^n(y)|$$

and in fact

$$|T^i(x) - T^i(y)| \leq \lambda^{-(n-i)} |T^n(x) - T^n(y)|.$$

The Folklore Theorem

Therefore we obtain

$$\begin{aligned} \log \frac{|DT^n(x)|}{|DT^n(y)|} &= K \sum_{i=1}^{n} \lambda^{-(n-i)} |T^n(x) - T^n(y)| \\ &\leq \frac{K}{\lambda - 1} |T^n(x) - T^n(y)| \leq \log K' \end{aligned}$$

for some $K' \in (1, \infty)$. This means that if $A \subset Z_n$ (so $T^n : A \to T^n(A)$ is a bijection), then

$$\frac{1}{K'} \frac{m(A)}{m(Z_n)} \leq \frac{m(T^n A)}{m(T^n Z_n)} = \frac{m(T^n A)}{m(\mathbb{S}^1)} \leq K' \frac{m(A)}{m(Z_n)}, \qquad (2)$$

where m is Lebesgue measure.

The Folklore Theorem

Construct the T-invariant measure μ. Take $B \subset \mathcal{B}$ arbitrary, and set $\mu_n(B) = \frac{1}{n}\sum_{i=0}^{n-1} m(T^{-i}B)$. Then by (2),

$$\frac{1}{K'}m(B) \leq \mu_n(B) \leq K'm(B).$$

We can take a weak* limit of the μ_n's; call it μ. Then

$$\frac{1}{K'}m(B) \leq \mu(B) \leq K'm(B),$$

and therefore μ and m are equivalent. The T-invariance of μ proven in the same way as in the Theorem of Krylov-Bogul'ubov.

The Folklore Theorem

Now for the ergodicity of μ, we need the Lebesgue Density Theorem, which says that if $m(A) > 0$, then for m-a.e. $x \in A$, the limit

$$\lim_{\varepsilon \to 0} \frac{m(A \cap B_\varepsilon(x))}{m(B_\varepsilon(x))} = 1,$$

where $B_\varepsilon(x)$ is the ε-balls around x. Points x with this property are called (Lebesgue) density points of A. (In fact, the above also holds, if $B_\varepsilon(x)$ is just a one-sided ε-neighbourhood of x.)

The Folklore Theorem

Assume by contradiction that μ is not ergodic. Take $A \in \mathcal{B}$ a T-invariant set such that $\mu(A) > 0$ and $\mu(A^c) > 0$. By equivalence of μ and m, also $\delta := m(A^c) > 0$. Let x be a density point of A, and Z_n be a neighbourhood of x such that $T^n : Z_n \to \mathbb{S}^1$ is a bijection. As $n \to \infty$, $Z_n \to \{x\}$, and therefore we can choose n so large (hence Z_n so small) that

$$\frac{m(A \cap Z_n)}{m(Z_n)} > 1 - \delta/K'.$$

Therefore $\frac{m(A^c \cap Z_n)}{m(Z_n)} < \delta/K'$, and using (2),

$$\frac{m(T^n(A^c \cap Z_n))}{m(T^n(Z_n))} \leq K' \frac{m(A^c \cap Z_n)}{m(Z_n)} < K'\delta/K' = \delta.$$

Since $T^n : A^c \cap Z_n \to A^c$ is a bijection, and $m(T^n Z_n) = m(\mathbb{S}^1) = 1$, we get $\delta = m(A^c) < \delta$, a contraction. Therefore μ is ergodic.