# El Teorema Ergódico

Pablo Lessa

9 de octubre de 2014

## 1. El Teorema Ergódico

Algún lector puede estar preguntándose ¿Qué tiene que ver la ley de grandes números con los temas tratados en la segunda sección?

La conexión es muy sencilla. Sea $\Omega = \{0,1\}^{\mathbb{N}}$ como en la primer sección. Consideramos la transformación $\sigma : \Omega \to \Omega$ definida a través de la ecuación

$$\sigma((\omega_1, \omega_2, \ldots)) = (\omega_2, \omega_3, \ldots).$$

La transformación $\sigma$ es conocida como 'el shift'. Para distinguirla de algunas variantes donde $\Omega$ es reemplazado por $\{0,1,\ldots,n-1\}^{\mathbb{N}}$ u otro conjunto de sucesiones, a veces se le denomina 'el shift de dos símbolos'.

El punto clave que coloca la ley de grandes números en el contexto de la teoría ergódica es que $\sigma$ preserva cualquiera de las probabilidades $\mu_p$ (que modelan el azar en el tiradas sucesivas de una moneda que tiene probabilidad $p$ de salir cara).

Con esta observación la ley de grandes números puede enunciarse de la siguiente forma

**Teorema 1** (Ley de Grandes Números)**.** *Para $\mu_p$-casi todo $\omega \in \Omega$ se cumple*

$$\lim_{n \to +\infty} \frac{f(\omega) + f(\sigma(\omega)) + \cdots + f(\sigma^{n-1}(\omega))}{n} = \mathbb{E}_{\mu_p}(f),$$

*donde $f(\omega) = \omega_1$ para $\omega = (\omega_1, \omega_2, \ldots)$.*

Por otro lado el Teorema de Khinchin puede enunciarse de la siguiente forma

**Teorema 2** (Khinchin)**.** *Sea $T : [0,1] \to [0,1]$ el mapa de Gauss, $\mu_T$ la medida invariante que cumple $\mu_T(A) = \frac{1}{\log(2)} \int_A 1/(1+x)\mathrm{d}x$ y $f : [0,1] \to \mathbb{N}$ definida por $f(x) = a_1$ donde $x = [0; a_1, a_2, \ldots]$. Entonces para $\mu_T$-casi todo $x \in [0,1]$ se cumple*

$$\lim_{n \to +\infty} \frac{\log(f(x)) + \log(f(T(x))) + \cdots + \log(f(T^{n-1}(x)))}{n} = \mathbb{E}_{\mu_T}\left(\log(f)\right).$$

**Ejercicio 1.** *Verificar que los dos enunciados que dimos del teorema de Khinchin son equivalentes.*

Hemos observado entonces que tanto la Ley de Grandes Números como el Teorema de Khinchin pueden enunciarse como afirmaciones sobre lo que le pasa al promedio de los valores de una función sobre segmentos de órbita $x, T(x), \ldots, T^n(x)$ cada vez más largos, para casi todo punto respecto a una probabilidad invariante para una cierta transformación $T$.

Agreguemos ahora un ejemplo más de este tipo (luego veremos otros). Para esto pido al lector que pause un poco para reflexionar sobre el siguiente ejercicio.

**Ejercicio 2.** *¿Cuál te parece que es el límite de la sucesión* $\frac{\sin(1)+\sin(2)+\cdots+\sin(n)}{n}$ *cuando* $n \to +\infty$*? ¿Podés dar una demostración?*

El siguiente teorema da otro ejemplo, a priori no relacionado, donde los métodos de la teoría ergódica pueden aplicarse.

**Teorema 3** (Equidistribución de Weyl)**.** *Sea* $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ *el círculo unidad en el plano complejo y* $\theta \in \mathbb{R}$ *un número irracional. Entonces la sucesión* $e^{2\pi i n\theta}$ *se equidistribuye en* $S^1$*, es decir para todo arco* $I \subset S^1$ *si* $1_I$ *es la indicatriz de* $I$ *se cumple*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_I(e^{2\pi i n\theta}) = m(I)$$

*donde* $m(I)$ *es la longitud del arco* $I$ *dividido el perímetro del círculo (i.e.* $2\pi$*).*

**Ejercicio 3** (Medida de Haar en $S^1$)**.** *Demostrar que la medida 'longitud de arco normalizada'* $m$ *en* $S^1$ *es la única probabilidad Boreliana que es invariante por toda rotación. Es decir, es la única probabilidad que cumple* $m(e^{it}A) = m(A)$ *para todo Boreliano* $A$*.*

## 1.1. Enunciado del teorema ergódico

La ley de grandes números, el teorema de Khinchin, y el teorema de equidistribución de Weyl pueden obtenerse a partir del siguiente resultado que es el punto de partida de la teoría ergódica.

**Teorema 4** (Teorema Ergódico)**.** *Sea* $X$ *un espacio métrico completo y separable,* $T : X \to X$ *una transformación medible,* $\mu$ *una probabilidad* $T$*-invariante, y* $f : X \to \mathbb{R}$ *una función medible y* $\mu$*-integrable. Entonces para* $\mu$*-casi todo* $x \in X$ *existe el límite*

$$\tilde{f}(x) = \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)).$$

*La función* $\tilde{f}$ *así definida cumple* $\mathbb{E}_\mu(f) = \mathbb{E}_\mu(\tilde{f})$*. Además si* $\mu$ *es una medida ergódica para* $T$ *entonces*

$$\tilde{f}(x) = \mathbb{E}_\mu(f)$$

*para* $\mu$*-casi todo* $x \in X$*.*

En la sección anterior verificamos que el mapa de Gauss $T$ era ergódico para la medida $\mu_T$. Por lo tanto el teorema de Khinchin se obtiene inmediatamente a partir del teorema ergódico. Otros resultados sobre los coeficientes de la fracción continua de los puntos típicos $x \in [0,1]$ son deducibles con el mismo método.

**Ejercicio 4.** *Demostrar que existe un número $\lambda_1$ tal que para Lebesgue-casi todo $x \in [0,1]$ el coeficiente $1$ aparece con frecuencia $\lambda_1$ en la fracción continua de $x$. Calcular el valor de $\lambda_1$.*

En principio por ahora sólo sabemos que $\mu_p$ es shift invariante pero no necesariamente ergódica en $\Omega = \{0,1\}^{\mathbb{N}}$. Por lo tanto sólo deduciríamos a partir del teorema anterior que el límite $\tilde{f}$ existe pero no que toma el valor $p$ en casi todo punto. Sin embargo la ergodicidad de $\mu_p$ para el shift es mucho más simple de demostrar que la ergodicidad del mapa de Gauss (y las ideas son similares).

**Ejercicio 5.** *Demostrar que para todo $p \in [0,1]$ la medida $\mu_p$ es ergódica para el shift de dos símbolos $\sigma$. Deducir la ley de grandes números a partir del teorema ergódico.*

La rotación $R_\theta(z) = e^{2\pi i\theta}z$ en $S^1$ con $\theta$ irracional tiene un propiedad sorprendente que lo diferencia por ejemplo del shift de dos símbolos. No solamente la medida de Haar $m$ es ergódica para $R_\theta$ sino que es **la única probabilidad invariante** para $R_\theta$.

**Ejercicio 6.** *Demostrar que si $\theta$ es irracional cualquier rotación $R_t$ es límite de cierta sucesión $R_{n_k\theta}$ donde $n_k \to +\infty$. Usando este hecho demostrar que $m$ es la única medida invariante para $R_\theta$. Deducir el teorema de equidistribución de Weyl a partir del teorema ergódico.*

**Ejercicio 7.** *Demostrar que para Lebesgue-casi todo $x \in [0,2\pi]$ el límite de $\frac{1}{n}\sum_{k=0}^{n-1}\sin(n+x)$ es $0$. Usando el hecho de que $\sin$ es $1$-Lipschitz deducir que de hecho el mismo resultado se cumple para $x=0$.*

## 1.2. Demostración del teorema ergódico: Caso ergódico

Lamentablemente no entiendo la demostración del teorema ergódico. No me mal interpreten, lo sé demostrar. El tema es que de la demostración, y a pesar de mucha reflexión sobre el tema, lo único que aprendo es que Birkhoff era muy buen analista. Se suele atribuir a David Hilbert la frase de que un resultado matemático no está realmente 'demostrado' hasta que no se pueda explicar porqué es verdad a una persona por la calle. En ese sentido el teorema ergódico sigue sin estar demostrado. Hay una larga genealogía de demostraciones, cada una de las cuales toma elementos de sus antecesores y 'simplifica' (según los autores) algún aspecto. Las últimas demostraciones de estos linajes son bastante cortas, pero todas tienen algún elemento que resta misterioso. Aquí les presento el argumento más sencillo que conozco para demostrar el teorema ergódico asumiendo que la transformación $T$ es ergódica para la medida $\mu$.

La demostración tiene un corazón combinatorio. Si $x_1, \ldots, x_n$ es una sucesión finita de números reales decimos que $k \in \{1, \ldots, n\}$ es un índice líder si existe $l \in \{k, k+1, \ldots, n\}$ tal que $x_k + \cdots + x_l \leq 0$. Como ejemplo, en la sucesión 2,1,-2,10,-8,0,4,-6 los índices líderes corresponden a los números 1,-2,10,-8,0,4,-6 (notemos que la suma es $-1$).

**Ejercicio 8** (Lema de los líderes (Riesz))**.** *Sea $x_1, \ldots, x_n$ una sucesión finita de números reales, entonces*

$$\sum_{k \text{ es un líder}} x_k \leq 0.$$

Aplicando el lema de los líderes a la sucesión $f(x), f(T(x)), f(T^2(x)), \ldots, f(T^n(x))$ se obtiene lo siguiente.

**Lema 1** (Karlsson y Margulis)**.** *Sea $X$ un espacio métrico separable y completo, $T : X \to X$ una transformación medible que preserva una probabilidad $\mu$, y $f : X \to \mathbb{R}$ una función medible con $\mathbb{E}_\mu(f) > 0$. Entonces*

$$\mu\left(\{x \in X : f(x) + \cdots + f(T^n(x)) > 0 \text{ para todo } n\}\right) > 0.$$

*Demostración.* Para todo par de naturales con $k < n$ definimos $A_{k,n}$ como el conjunto de los $x \in X$ tales que el índice $k$ es líder para la sucesión $f(x), f(T(x)), \ldots, f(T^n(x))$. Por el lema de Riesz se cumple

$$\mathbb{E}_\mu\left(f(x)1_{A_{1,n}} + \cdots + f(T^n(x))1_{A_{n,n}}\right) \leq 0.$$

Como $T$ preserva $\mu$ la integral de $g \circ T$ coincide con la de $g$ para toda $g : X \to \mathbb{R}$ integrable. Aplicando esto a la desigualdad anterior se obtiene

$$\mathbb{E}_\mu\left(f(x)\left(1_{B_n} + 1_{B_{n-1}} + \cdots + 1_{B_1}\right)\right) \leq 0,$$

donde $B_k$ es el conjunto de los $x \in X$ tales que $f(x) + \cdots + f(T^l(x)) \leq 0$ para algún $l \leq n$.

Dividiendo entre $n$ y aplicando el teorema de convergencia dominada se obtiene

$$\mathbb{E}_\mu\left(f(x) \limsup_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} 1_{B_k}\right) \leq 0$$

Como $\mathbb{E}_\mu(f) > 0$ el conjunto de los $x$ tales que $\limsup\limits_{n\to+\infty} \frac{1}{k}\sum 1_{B_k} = 1$ no puede tener medida total. Esto implica que un conjunto de medida positiva de $x \in X$ pertenece a infinitos $B_k$, lo cual implica el enunciado. □

El lema de Karlsson y Margulis permite dar una demostración directa del Teorema Ergódico en el caso ergódico.

*Demostración del teorema ergódico: Caso ergódico.* Asumimos que $\mathbb{E}_\mu(f) = 0$ (el caso general se deduce de este caso aplicándolo a $f - \mathbb{E}_\mu(f)$).

Aplicando el lema de Karlsson y Margulis a $f+\epsilon$ se obtiene que el conjunto de los $x$ tales que

$$f(x)+\cdots+f(T^n(x))+n\epsilon \geq 0$$

para todo $n$ es de $\mu$-medida positiva. Esto implica que el conjunto

$$A_\epsilon = \left\{ x \in X : \limsup_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x)) \leq \epsilon \right\}$$

cumple $\mu(A_\epsilon) > 0$.

Como $\mu$ es ergódica y $A_\epsilon$ es invariante se tiene $\mu(A_\epsilon) = 1$. Como esto se cumple para todo $\epsilon > 0$ (en particular los racionales) hemos demostrado que

$$\limsup_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x)) \leq 0$$

para $\mu$-casi todo $x$.

El mismo argumento aplicado a $-f$ muestra que el $\liminf$ es mayor o igual a 0 para $\mu$-casi todo $x$. Con esto se concluye el teorema. $\square$

## 1.3. Demostración del teorema ergódico: Caso general

La diferencia fundamental entre el caso ergódico y el caso no ergódico se pueden dar ejemplos en los cuales el límite

$$\tilde{f}(x) = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$$

toma diferentes valores con probabilidad positiva.

**Ejercicio 9.** *Comprobar la afirmación anterior (Sugerencia: Toda permutación de un conjunto finito preserva la medida de conteo normalizada, ¿Qué comportamiento puede tener $\tilde{f}$ en esta familia de ejemplos?).*

Para reducir la demostración al caso en el cual queremos mostrar que $\tilde{f} = 0$, en el caso ergódico alcanzaba con cambiar $f$ por $f - \mathbb{E}_\mu(f)$. En el caso no ergódico hay que restarle a $f$ una función, llamada la esperanza condicional de $f$ a la $\sigma$-álgebra de conjuntos invariantes, un poco más difícil de identificar.

**Ejercicio 10.** *Sea $T : X \to X$ una transformación que preserva una probabilidad $\mu$ en un espacio métrico separable y completo $X$. Denotemos por $\mathcal{F}^{inv}$ la familia de Borelianos que son $T$-invariantes.*

1. *Demostrar que $\mathcal{F}^{inv}$ es una $\sigma$-álgebra.*
2. *Demostrar que para toda $f : X \to \mathbb{R}$ que es $\mu$-integrable existe una función $g : X \to \mathbb{R}$ (en notación probabilística $g = \mathbb{E}_\mu(f|\mathcal{F}^i nv)$) que es $T$-invariante y $\mu$-integrable y que satisface $\int_A f(x)\mathrm{d}\mu(x) = \int_A g(x)\mathrm{d}\mu(x)$ para todo Boreliano invariante $A \in \mathcal{F}^{inv}$ (Sugerencia: Usar el teorema de Radon-Nikodym para las medidas $\mu$ y $A \mapsto \mathbb{E}_\mu(f 1_A)$ en $\mathcal{F}^{inv}$).*

Con esto podemos demostrar el teorema ergódico en el caso general.

*Demostración del teorema ergódico: Caso general.* Podemos asumir que $f$ integra 0 en todo Boreliano invariante (el caso general se reduce a este cambiando $f$ por $f - \mathbb{E}_\mu(f|\mathcal{F}^{\text{inv}})$).

Por el lema de Karlsson y Margulis dado $\epsilon > 0$ el conjunto

$$A_\epsilon = \left\{ x : \limsup \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) \leq \epsilon \right\}$$

tiene medida positiva.

Sea $B_\epsilon = X \setminus A_\epsilon$. Si $\mu(B_\epsilon) > 0$ entonces aplicando el lema de Karlsson y Margulis a la medida que cumple $\tilde{\mu}(A) = \mu(A \cap B_\epsilon)/\mu(B_\epsilon)$ para todo $A$, obtenemos que $A_\epsilon$ también tiene medida para positiva para $\tilde{\mu}$ lo cual es absurdo. Por lo tanto $\mu(A_\epsilon) = 1$.

Como esto es valido para todo $\epsilon > 0$ (y también repitiendo el argumento para $-f$ se obtiene el resultado. $\square$