EXPERIMENTOS CON SISTEMAS DINÁMICOS CAÓTICOS PARA SIMULAR EN COMPUTADORA

Eleonora Catsigeras

Propuesta de colaboración para el curso de Matlab de Claudia Stocco
Noviembre de 2002.

1. DEFINICIÓN DE T^2 (Toro bidimensional).

Consideremos el conjunto siguiente:

$$\boldsymbol{T}^2 = \{(x, y) \in \mathbb{R}^2\}_{\text{mod. }(1,1)}$$

donde mod. $(1, 1)$ significa que se identifican todas las parejas (x, y) que difieren de una dada en un vector de componentes enteras. Esto es:

(x_1, y_1) es equivalente a (x_2, y_2) o

$$(x_1, y_1) =_{\text{mod. }(1,1)} (x_2, y_2)$$

si, por definición: $x_1 - x_2$ e $y_1 - y_2$ son números enteros.

Cada clase de equivalencia (un subconjunto de $\mathbb{R}^2$ lo más grande posible, formado por parejas que son todas equivalentes entre sí)tiene un único representante $(x, y) \in \mathbb{R}^2$ tal que $0 \leq x < 1$ y $0 \leq y < 1$. Las demás parejas de la misma clase se obtienen de ese representante sumando un vector cualquiera de coordenadas enteras.

T^2 es un toro (bidimensional) porque se puede interpretar como el cuadrado $[0, 1] \times [0, 1] = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1; 0 \leq y \leq 1\}$ de $\mathbb{R}^2$ en el que el lado $x = 1$ se identifica (se pega) con el lado $x = 0$, y el lado $y = 1$ se identifica (se pega) con el lado $y = 0$.

En T^2 definimos la siguiente métrica:

$$dist((x_1, y_1), (x_2, y_2))_{\text{mod. }(1,1)} = \sqrt{\text{mín}\,\{|x_1 - x_2 + i|^2 + |y_1 - y_2 + i|^2 : i = 0, 1, -1\}}$$

Observar que la métrica definida es la usual en el cuadrado de lado 1 de $\mathbb{R}^2$, si la diferencia de abscisas (y la de ordenadas) de los puntos es menor o igual que 0.5 en valor absoluto. En cambio si la diferencia de abscisas (o la de ordenadas) es mayor que 0.5 (por ejemplo $x_1 = 0{,}9,\ y_1 = 0,\ x_2 = 0,2,\ y_2 = 0,\ x_1 - x_2 = 0{,}7$), la distancia usual en el cuadrado es mayor que 0.5 (en el ejemplo, es igual a 0.7). Sin embargo esta última no es la distancia en el toro que resulta de identificar el lado $x = 0$ con el lado $x = 1$. En el toro la "diferencia de abscisas "es la mínima diferencia de abscisas entre todos los puntos del plano equivalentes módulo $(1, 1)$ con (x_1, y_1) y todos los puntos del plano equivalentes módulo $(1, 1)$ con (x_2, y_2). En el ejemplo, queda: $(1 + x_2) - x_1 = 1{,}2 - 0{,}9 = 0{,}3$

Con esta estructura de espacio métrico, se definen las propiedades topológicas de $\boldsymbol{T}^2$ (abiertos, entornos, compacidad, etc), y las funciones continuas $g : \boldsymbol{T}^2 \mapsto \mathbb{R}$.

Resulta que g es una función $\boldsymbol{T}^2 \mapsto I\!\!R$ si y solo si es "el cociente"(la restricción al cuadrado de lado 1 en R^2) de una función $\tilde{g} : I\!\!R^2 \mapsto I\!\!R$, llamada levantamiento de g a $I\!\!R^2$, que cumple $g(x+n, y+m) = g(x,y)$ para todo $(x,y) \in I\!\!R^2$ y para todos m,n enteros. Además g es continua si y solo sí su levantamiento $\tilde{g}$ es continua.

De manera similar $\boldsymbol{T}^2$ hereda de $I\!\!R^2$ la estructura diferenciable (variedad diferenciable). Así $g : \boldsymbol{T}^2 \mapsto I\!\!R$ es diferenciable hasta orden k con derivadas continuas hasta orden k (en breve $g \in C^k$) si y solo si su levantamiento $\tilde{g}$ lo es en $I\!\!R^2$.

Definimos el espacio medible (espacio probabilizable)

$$\boldsymbol{T}^2 \times \mathcal{B}$$

donde $\mathcal{B}$ indica la sigma- álgebra de Borel.

En ese espacio medible interesan **todas las probabilidades**, y en particular encontrar una probabilidad adecuada a la dinámica que vamos a definir más adelante.

Sin embargo, cuando no se dice lo contrario, vamos a considerar como probabilidad p privilegiada, la uniformemente distribuida en $\boldsymbol{T}^2$, es decir, la que se hereda del área de los abiertos en $I\!\!R^2$. Así, dado un abierto $V \subset \boldsymbol{T}^2$, por ejemplo el representado por un único cuadradito de lado menor que $1/2$ en $I\!\!R^2$, la probabilidad de V será el área de ese cuadradito. Esa probabilidad, extendida a todos los Borelianos y que vale cero en todos los subconjuntos contenidos a algún Boreliano de probabilidad cero, se llama "medida de Lebesgue"ó abusando del lenguaje: "área".

EXPERIMENTO 1: Generar puntos P de T^2 aleatoriamente con la distribución de probabilidad uniforme "área"p. Elegir cuadraditos Q de lado $\epsilon > 0$ en $\boldsymbol{T}^2$ ($\epsilon < 1/2$) y diferentes centros (por ejemplo centros (0,0), (1/2, 1/2), (1,0), etc) y hacer un programa que genere puntos P aleatoriamente con distribución uniforme y cuente los puntos que cayeron en cada cuadradito Q. Calcular la relación entre esa cantidad y la cantidad N de puntos que se generaron (frecuencia de visita a cada Q). Comparar la frecuencia de visita a Q con el área del cuadradito (ϵ^2), cuando N aumenta. Después variar ϵ acercándolo a cero. ¿ Qué sucede si ϵ se acerca mucho a cero? ¿Cuánto hay que esperar (es decir: qué valor de N en función de ϵ) para que la frecuencia de visita a Q no difiera del área más que $1\,\%$ del área de Q.

2. TRANSFORMACIÓN QUE PRESERVA LA PROBABILIDAD. MEDIDAS DE PROBABILIDAD INVARIANTES POR F.

Consideremos el siguiente ejemplo de transformación continua $F : \boldsymbol{T}^2 \mapsto \boldsymbol{T}^2$ del espacio.

He aquí un ejemplo de transformación "lineal"en T^2:

$$F((x,y)_{\text{mod. }(1,1)} = (2x+y, x+y))_{\text{mod. }(1,1)}$$

que tiene como matriz asociada $[2,1;1,1]$.

Esa transformación F está bien definida del toro $\boldsymbol{T}^2$ a sí mismo, porque al sumar a la pareja (x,y) de reales una pareja de enteros (m,n) cualquiera, resulta la pareja $(2x+y, x+y)$ de R^2 también incrementada en cantidades enteras.

Se puede demostrar fácilmente que F es invertible (ojo: en el toro), y hallar su función inversa F^{-1}.

F es además diferenciable, de clase C^k para todo k natural (es decir $F \in C^\infty$).

Como el determinante de la matriz derivada de F es igual a 1, se demuestra que F preserva el área, o sea la probabilidad p que llamamos medida de Lebesgue o área. Es decir:

$$p(F^{-1}(B)) = p(B) \text{ para todo } B \in \mathcal{B} \quad \text{(ver nota más abajo)}$$

En particular alcanza con demostrar lo anterior poniendo en lugar de B todo cuadradito de lado $\epsilon > 0$, para todo ϵ.

NOTA: $F^{-1}(B)$ indica preimagen del conjunto B por F y existe, aunque sea vacía, para cualquier F y para cualquier B, aunque F no sea invertible. Usamos la notación $p(B)$ para indicar la probabilidad de que un punto P generado aleatoriamente con distribución de probabilidad p caiga en el conjunto B.

EXPERIMENTO 2: Inventar un programita de computadora para verificar lo mejor que se pueda que la transformación F del ejemplo dado al principio preserva la probabilidad p uniformemente distribuida (el área).

EXPERIMENTO 3: Sea ahora la transformación $G : \boldsymbol{T}^2 \mapsto G$ definida por

$G((x,y)_{\text{mod }(1,1)} = \big(2x+y-\frac{1}{2\pi}[\text{sen}(2\pi x)\cos[2(\pi y)], x+y-\frac{1}{2\pi}[\text{sen}(2\pi x)\cos[2(\pi y)]\big)_{\text{m}}$

Verificar con la computadora lo más que se pueda que G NO preserva la probabilidad p uniformemente distribuida (el área).

TEOREMA 1 (Existencia de probabilidades invariantes)

Dada cualquier transformaci'on continua F de un espacio métrico compacto en sí mismo, existe alguna probabilidad de Borel p que es invariante por F, o sea, que es preservada por F. Más precisamente: $(p(F^{-1}(B)) = p(B)$ para todo $B \in \mathcal{B}$.

Comentarios sobre el Teorema 1:

Usualmente existen infinitas probabilidades p invariantes para una misma F dada. Esto sucede en los ejemplos de las transformaciones F y G usadas en los experimentos anteriores.

Usualmente ninguna de esas probabilidades invariantes es absolutamente continua respecto al área. (Es decir no existe una función densidad en L_1 que permita calcular la probabilidad de cualquier abierto V integrando la función densidad en el abierto respecto al área). Sin embargo en el caso de la transformación F dada como ejemplo más arriba, existe una única probabilidad invariante que es absolutamente continua respecto al área p (y es justamente el área).

CONJETURA 1

Para la transformación G dada como ejemplo en el experimento 3, existe una única probabilidad invariante que es absolutamente continua respecto al área.

(La demostración matemática de la unicidad es fácil; lo complicado es la existencia, y hoy en día no se conoce si la afirmación de existencia de esta conjetura es verdadera o falsa).

Frecuentemente pasa que todas las infinitas probablidades invariantes para una misma transformación F dada (que no preserva el área) son singulares respecto al área p.(Es decir están concentradas en borelianos con área cero).

La demostración del teorema 1 da un procedimiento para construir alguna(s) medidas invariantes para una transformación F continua dada del espacio.

IDEA DE LA DEMOSTRACIÓN DEL TEOREMA 1.

Consiste en tomar inicialmente cualquier probabilidad p_1 en el espacio medible de Borel. La probabilidad p_1 no será necesariamente invariante por F: por ejemplo p_1 es cualquier delta de Dirac que uno elija.

Luego uno define una sucesión de probabilidades, (usualmente ninguna queda invariante), así:

$$p_n = \frac{1}{n}\sum_{j=1}^{n-1} {F^j}^* p_1 \quad \text{(ver nota más abajo)}$$

Habrá que demostrar que existe una subsucesión $\{p_{n_i}\}_{i\in\mathbb{N}}$ que converge en la topología débil * (ver nota más abajo), a un límite p. Este límite p resulta ser una probabilidad invariante por F, cuando F es continua.

Nota: Dada $F : X \mapsto X$ donde X es un espacio métrico compacto, el conjunto $\mathcal{P}$ de todas las probabilidades de Borel en X queda sometido a la transformación $F^* : \mathcal{P} \mapsto \mathcal{P}$ del siguiente modo:

$$F^*(p)(B) = p(F^{-1}(B))$$

En $\mathcal{P}$ se define la topología débil * así:

$$\overset{w^*}{\lim_{i\to\infty}} p_i = p$$

si

$$\lim_{i\to\infty} \int g\, dp_i = \int g\, dp$$

para toda $g : X \mapsto \mathbb{R}$ continua. Si X es un espacio métrico compacto, el espacio $\mathcal{P}$ con la topología w^* es compacto y metrizable, y por lo tanto secuencialmente compacto.

EXPERIMENTO 4: Inventar un programa con la computadora para visualizar (en forma aproximada) alguna probabilidad $\hat{p}$ en $(\boldsymbol{T}^2, \mathcal{B})$ que sea preservada por la transformación G del experimento 3.

Si se consiguiera una probabilidad $\hat{p}$ invariante por G y que además (¿cómo saberlo?) fuera absolutamente continua respecto al área (suponiendo que la afirmación de la conjetura 1 fuera verdadera), graficar aproximadamente su densidad.

La densidad debería ser una función real mayor que cero definida en el toro bidimensional. El toro puede representarse como el cuadrado de lado 1 en $\mathbb{R}^2$. De esa forma se podría hacer una gráfica tridimensional de la función densidad, o bidimensional pintando con colores de tonos más claros los puntos del cuadrado donde la densidad es mayor y negros donde es cero.

3. DINÁMICA DE ITERADOS DE LA TRANSFORMACIÓN F, SISTEMAS CAÓTICOS Y TOPOLÓGICAMENTE ESTABLES.

Sea $F : \boldsymbol{T}^2 \mapsto T^2$ continua cualquiera.

Sea P_0 un punto cualquiera de $\boldsymbol{T}^2$. Llamaremos a P_0 como *DATO INICIAL* , o estado inicial, o punto inicial.

DEFINICIÓN: Se llama *ÓRBITA (o trayectoria) por P_0 según F* a la sucesión de puntos $P_0, P_1 = F(P_0), \ldots, P_n = F(P_{n-1}(P_0) = F^n(P_0), \ldots$ para todo $n \geq 1$ natural. Aquí F^n indica composición de F consigo misma n veces.

Cuando F es invertible, se llama ÓRBITA (o trayectoria) negativa por P_0 según F a la órbita positiva según F^{-1}. Es decir, la sucesión: $P_0, \ldots, P_{-n} = F^{-1}(P_{-(n-1)} = F^{-n}(P_0), \ldots$ para todo $n \geq 1$ natural. Aquí F^{-n} indica composición de F^{-1} consigo misma n veces.

Cuando F es invertible se llama ÓRBITA (o trayectoria) completa por P_0, a la sucesión biinfinita $\{P_n = F^n(P_0)\}_{n \in Z}$ para n entero. (Aquí $n \in Z$ significa que el índice de la sucesión varía en todos los enteros).

DEFINICIONES: La *dinámica (o sistema dinámico) por iterados de F* es la transformación que a cada punto P_0 del espacio le hace corresponder su ÓRBITA COMPLETA $\{P_n\}_{n \in Z}$ por F.

La dinámica es *expansiva o caótica* (o, en breve, la transformación F es expansiva o caótica) si existe una constante de ERROR inevitable (constante de expansividad) $\alpha > 0$ tal que PARA TODOS LOS DATOS INICIALES P_0 y Q_0, se cumple:

$$\text{dist}\,(P_n, Q_n) > \alpha \text{ para algún } n \in Z$$

Significado: aunque inicialmente P_0 y Q_0 sean muy próximos, porque se comete un error inicial arbitrariamente pequeño, (pero positivo), el error al encontrar el punto n-ésimo de la órbita por P_0 y por Q_0 se hace necesariamente mayor que una constante $\alpha > 0$ (α depende solo de la transformación F dada), para cierto instante n (que depende de P_0 y Q_0). Por lo tanto no podemos darnos el lujo de alterar en lo más mínimo, ni aproximar nada el dato inicial P_0 (ni sus futuras posiciones al aplicar reiteradamente la transformación F), porque si lo hacemos, en vez de conocer la órbita del punto P_0 verdadero, veremos una órbita que en algún momento se separa groseramente (más que $\alpha > 0$) de la verdadera.

TEOREMA 2. *Las transformaciones F y G definidas en los dos ejemplos anteriores, son caóticas.*

EXPERIMENTO 5 Tomar las transformaciones F y G de los experimentos anteriores y calcular a mano exactamente los primeros 10 iterados de los puntos iniciales $(1/2, 1/2)$ y $(1/3, 1/3)$. Hacer un programita para que la computadora encuentre las órbitas respectivas y comparar los resultados con los calculados exactamente.

Sin embargo, la presencia del caos o la expansividad no es tan desalentadora como parece, ni es causa, como (creo que erróneamente se dice en algunos cursos de métodos numéricos), para que "no sean válidos los experimentos con computadora para estudiar las transformaciones caóticas o expansivas". Es solo una advertencia de que hay que estudiar con cuidado qué es lo que uno ve en la computadora, antes de creer que lo que ve es exactamente o aproximadamente lo que uno se propuso ver.

La razón de que las transformaciones F y G de los ejemplos dados al principio puedan ser sí estudiadas en la computadora a pesar de ser caóticas con tal de que se sepan interpretar los resultados, es que uno puede estudiar la dinámica en sentido global (y no particular órbita por órbita), debido a lo siguiente:

DEFINICIÓN Una transformación continua F en un espacio métrico compacto X es *topológicamente estable* , si tiene la siguiente propiedad:

PROPIEDAD DE ACOMPAÑAMIENTO DE LAS PSEUDO-ÓRBITAS POR ÓRBITAS VERDADERAS.

Dado $\epsilon > 0$ existe $\delta > 0$ tal que, para toda sucesión $\{Q_n\}_{n \in Z}$ (llamada δ- pseudo-órbita), que cumpla:

$$dist(F(Q_n), Q_{n+1}) < \delta \quad para\ todo\ n \in Z$$

EXISTE UNA ÚNICA ÓRBITA COMPLETA $\{\tilde{P}_n\}_{n \in Z}$ por F (verdadera órbita por F), que cumple:

$$dist(\tilde{P}_n, Q_n) < \epsilon \ para\ todo\ n \in Z$$

Explicación: Está especificado un error máximo tolerable *epsilon* > 0 entre los resultados a obtener en la computadora y los reales. Si F es topológicamente estable, existe un $\delta > 0$, que es la alteración máxima de redondeo que deberá hacer la computadora: ya sea al ingresar el dato inicial Q_0 en vez del verdadero dato inicial que yo tenía P_0, como al calcular los sucesivos Q_n redondeando el valor de $F(Q_{n-1})$.

La sucesión de puntos observada es la de los puntos Q_n que difiere groseramente de la órbita verdadera del dato inicial que yo tenía P_0.

Sin embargo, lo que veo en la computadora **es la ϵ-aproximación en todo n, (con ϵ arbitrariamente chico especificado al principio) de CIERTA ORBITA VERDADERA CON UN PUNTO INICIAL $\tilde{P}_0$, desconocido pero existente, y que está además $\delta + \epsilon$ cercano a P_0.**

TEOREMA 3 Las transformaciones F y G definidas como ejemplos hasta ahora son topológicamente estables, o sea tienen la propiedad de acompañamiento de las pseudo-órbitas por órbitas verdaderas.

PROBLEMA 6: Volver al experimento 4, y analizar si la probabilidad $\hat{p}$ visualizada aproximadamente en ese experimento tiene alguna relación de proximidad con alguna probabilidad $\tilde{p}$ verdaderamente invariante por G.

PROBLEMA 7: Problema molesto si los hay, y que depende además de cómo la computadora aproxime o redondee.

Considérese la transformación F o la G de los experimentos anteriores. Sea P_0 un punto que se elige aleatoriamente con distribución de probabilidad uniforme en el toro T^2. Sea $\{Q_n\}_{n\in Z}$ LA δ-pseudo-órbita (y no UNA cualquiera sino LA única) que se visualizaría en la computadora para todo $n \in Z$ cuando se ingresa como dato inicial P_0. Sea $\tilde{P}_0$ el único punto inicial (desconocido pero existente) de la órbita verdadera que acompaña a esa pseudo-órbita. Se observa que $\tilde{P}_0$ es una función desconocida pero existente de P_0 ¿Cuál es la distribución de probabilidad de $tildeP_0$?

DISGRESIÓN PERSONAL RESPECTO AL PROBLEMA 7:

La computadora tiene una cantidad finita de memoria y cuando redondea en realidad discretiza, proyecta a un conjunto finito. Por otro lado supongo que el programa está hecho para que en cada paso n de iteración, ante un mismo dato Q_n produzca la computadora el mismo dato de salida Q_{n+1}. Entonces todas las pseudo-órbitas son periódicas, y son una cantidad finita. Por lo tanto hay solo una cantidad finita de puntos $\tilde{P}_0$ posibles. Es decir la distribución de probabilidad de los verdaderos puntos iniciales de las únicas verdaderas órbitas que son visibles en la computadora es solo una combinación de deltas de Dirac (seguramente muchísimas deltas de Dirac, pero una cantidad finita). Por lo tanto, lo que uno ve es el comportamiento, muy próximo a órbitas reales sí, y de muchísimas órbitas, pero solo de una cantidad finita. Conclusión: al experimentar con la computadora "minga"de que los conjuntos de puntos iniciales de las órbitas observables o relevantes, o probables, o abundantes, o como las quieran llamar, son los conjuntos con área (o medida de Lebesgue) positiva.

4. FRECUENCIAS DE VISITA DE LAS ÓRBITAS.

Sea F una transformación continua de un espacio métrico en sí mismo. Sea P_0 un punto inicial. Y sea B un subconjunto boreliano del espacio.

Definición: Se llama frecuencia de visita de P_0 a B, o promedio temporal de estadía de P_0 en B hasta el instante $N \geq 1$ al número racional no negativo ni mayor que uno definido por:

$$T_N(P_0, B) = \frac{\#\{j \text{ número natural} : 0 \leq j \leq N-1; P_j \in B\}}{N}$$

donde $P_j = F^j(P_0) = F(P_{j-1})$ para los naturales $j \geq 1$

Se llama frecuencia asintótica de visita de P_0 a B, o promedio temporal de estadía asintótico de P_0 en B, cuando existe, al número real $T(P_0, B)$ no negativo ni mayor que uno igual al límite de $T_N(P_0, B)$ cuando N tiende a infinito, si existe ese límite.

El siguiente resultado es el fundamento de la teoría ergódica de los sistemas dinámicos:

TEOREMA 4 (Teorema de Birkhoff y teorema de existencia de medidas ergódicas)

Para toda probabilidad p invariante por F, para todo B boreliano, el tiempo $T(P_0, B)$ existe p-c.t.p. P_0.

Además $T(\cdot, B)$ es una función $L_1(p)$ y su integral respecto a la probabilidad p es igual a $p(B) = p(X_0 \in B)$.

Además cualquiera sea F continua en un espacio métrico compacto X, existe(n) probabilidades de Borel p en X que son F- invariantes y para las cuales $T(P_0, B)$ es constante p-c.t.p. P_0 e igual a $p(B)$.

Definición Esas últimas probabilides se llaman *ergódicas.* Para ellas los promedios temporales asintóticos de estadía de P_0 en el conjunto B (que son constantes p casi todo punto P_0 pero que dependen de B) son iguales al "promedio espacial"del conjunto B, es decir a $p(B)$, o valor esperado de la función característica χ_B del conjunto B.