Jest to tekst związany z odczytem wygłoszonym na XLVIII Szkole Matematyki Poglądowej, *Skojarzenia i analogie*, Otwock–Śródborów, styczeń 2012.

Układy dynamiczne i rachunek prawdopodobieństwa

Michał SZOSTAKIEWICZ, Warszawa

Układy dynamiczne to (pół)grupy przekształceń na przestrzeni (topologicznej, metrycznej, mierzalnej…) X. W naszym przypadku będzie to półgrupa generowana przez jedno przekształcenie $T : X \to X$ (czyli izomorficzna z $\mathbb{N}$). Zadaniem teorii układów dynamicznych jest opis *trajektorii* (inna nazwa: *orbit*) punktów $x \in X$, to znaczy ciągów $(x, T(x), T^2(x), \ldots)$, gdzie T^i oznacza i-krotne złożenie przekształcenia T.

Okazuje się (teoria chaosu), że opisanie *każdej* trajektorii może być praktycznie niemożliwe, ale *teoria ergodyczna* daje możliwość powiedzenia czegoś o *prawie każdej* trajektorii względem pewnej miary (probabilistycznej).

Inna motywacja do zajmowania się *teorią ergodyczną* ma związek z rachunkiem prawdopodobieństwa (i na nim się skupimy). Przypomnijmy, że metoda Monte Carlo mówi, że jeśli wylosujemy ciąg punktów $x_0, x_1, \ldots$ ze zbioru X niezależnie, z pewnym rozkładem μ, zaś $\phi : X \to \mathbb{R}$ jest funkcją całkowalną, to ciąg $\frac{1}{n}\sum_{i=0}^{n-1} \phi(x_i)$ zbiega prawie na pewno do $\int \phi\, d\mu$. Wynika to po prostu z MPWL (Mocnego Prawa Wielkich Liczb) Kołmogorowa. Możemy tę zbieżność interpretować tak, że ciąg x_i „równomiernie" wypełnia przestrzeń X. Będziemy chcieli pozbyć się losowości i jako ciąg x_i wybrać trajektorię $T^i(x)$ dla pewnych T i $x \in X$.

Rozpatrzmy na początek dwa przykłady.

Weźmy $X = S^1$ okrąg jednostkowy i przez $R_\alpha : S^1 \to S^1$ oznaczmy obrót okręgu o kąt $2\pi\alpha$ dla $\alpha \in [0, 1]$. Powszechnie wiadomo, że jeśli α jest niewymierna, to trajektoria dowolnego punktu na okręgu jest gęsta. Można ponadto pokazać (autorowi nie jest znany dowód elementarny), że ta trajektoria jest „równomiernie rozłożona" w następującym sensie:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(R_\alpha^i(x)) = \int \phi\, d\mathrm{Leb} \qquad (*)$$

gdzie Leb to (unormowana) miara Lebesgue'a, zaś ϕ jest (prawie) dowolną funkcją całkowalną. Związek z Monte Carlo jest ewidentny.

Należy podkreślić, że nie jest to tylko mniej lub bardziej interesująca ciekawostka, ale fakt z wieloma konsekwencjami. Jedna z nich wynika z interpretacji wyrażenia $\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A(R_\alpha^i(x))$ jako częstości wpadania trajektorii punktu x do zbioru $A \subseteq X$ do czasu n.[1] W związku z tym z $(*)$ wynika, że ta częstość asymptotycznie jest równa mierze zbioru A. To z kolei pozwala rozwiązać następujący problem[2]:

Zadanie 1. Rozpatrzmy ciąg pierwszych cyfr zapisu dziesiętnego liczb postaci 2^n (to znaczy $1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, 8, 1, 3, 6, 1, 2, 5, \ldots$). Czy w tym ciągu częściej (asymptotycznie) będzie się pojawiała cyfra 7 czy 8?

Zadanie jest o tyle ciekawe, że o ile 8 widzimy na czwartej pozycji, to na pierwszy rzut oka nie jest jasne, czy cyfra 7 w ogóle się pojawi (tym bardziej, że pojawia się za pierwszym razem dopiero na 46. miejscu). Równie nieoczywiste może być też to, czy ten ciąg jest, czy nie jest okresowy.

Wszystkie te pytania stają się bardzo proste, kiedy rozpatrzymy części ułamkowe liczb $\log 2^n = n \log 2$ (tym razem log to logarytm dziesiętny). Z jednej strony tworzą one dokładnie trajektorię punktu 0 przy obrocie $R_{\log 2}$ (i utożsamieniu S^1 z $[0, 1)$). Z drugiej, każdy Czytelnik łatwo napisze warunek na to, że pierwszą cyfrą dziesiętną liczby N jest c, zależący tylko od części ułamkowej $\log N$. Ten warunek (i zbieżność $(*)$) implikuje, że cyfra 7 będzie się pojawiała częściej niż 8.

---

[1] wynika ona oczywiście z równości $\sum_{i=0}^{n-1} \mathbb{1}_A(R_{\alpha^i(x)}) = \#\{0 \leqslant i < n : R_\alpha^i(x) \in A\}$

[2] por. np. artykuły Pawła Strzeleckiego w *Delcie* z sierpnia 1994 lub marca 2010

Zbieżność $(*)$ nie jest właściwie niczym nadzwyczajnym, lecz konsekwencją Twierdzenia Ergodycznego Birkhoffa, o którym będzie mowa później. Zanim je sformułujemy, podamy (trochę) bardziej skomplikowany przykład przekształcenia T, a wcześniej poczynimy ważną uwagę.

Otóż miara Lebesgue'a, która pojawia się we wzorze $(*)$ ma ważną i nieprzypadkową własność: jest niezmiennicza dla przekształcenia $T = R_\alpha$ w takim sensie, że dla dowolnego zbioru (mierzalnego) $A \subseteq S^1$ zachodzi równość $\mu(T^{-1}(A)) = \mu(A)$ (w sposób zwarty — i mądry — możemy to zapisać tak: $T_*\mu = \mu$). Własność ta jest kluczowa dla Twierdzenia Ergodycznego, a ponadto ma swoją naturalną „fizykalną" interpretację jako stan stacjonarny naszego układu.

Pora na obiecany drugi przykład: dwukrotne nawinięcie okręgu. Przestrzenią znowu jest $X = S^1$ a przekształcenie — po utożsamieniu z S^1 z $[0,1)$ — zadane wzorem $T(\theta) = 2\theta \bmod 1$ (możemy też w sposób zwarty zapisać to multiplikatywnie w zmiennych zespolonych jako z^2 dla $z \in \mathbb{C}$ t.że $|z| = 1$).

Ten przykład wygląda być może niepozornie, ale jego dynamika jest dużo bardziej skomplikowana. W szczególności poprzednio każde dwie trajektorie były właściwie takie same, to znaczy różniły się od siebie tylko o pewien obrót. Teraz ich zachowanie może być diametralnie różne, w szczególności na pewno nie każda orbita jest równomiernie rozłożona na okręgu.

Najbardziej ekstremalnym tego przykładem są punkty okresowe (to znaczy spełniające $T^p(z) = z$ dla pewnego $z \in S^1$) — takich punktów dla obrotu (o kąt niewymierny!) nie było. W takim wypadku ciąg $\frac{1}{n}\sum_{i=0}^{n-1} \phi(T^i(z))$ zbiega do średniej wartości ϕ na trajektorii z. Punkty okresowe są — w pewnym sensie — przeciwieństwem punktów o gęstej trajektorii, bo odwiedzają tylko „bardzo mały fragment" zbioru X.

Jako dobre ćwiczenie można się zastanowić, dlaczego dla przekształcenia T istnieją punkty okresowe o dowolnym okresie (podstawowym).

Pokazanie, że istnieje punkt o orbicie gęstej jest nie-elementarne w tym sensie, że autorowi nie jest znany dowód nie korzystający z przestrzeni symbolicznej (to znaczy przestrzeni nieskończonych słów nad pewnym alfabetem). Można jednak pokazać, że zbiór takich punktów jest ich bardzo duży, a konkretnie, że nawet zbiór punktów o trajektoriach równomiernie rozłożonych (który jest jego podzbiorem), tworzy zbiór pełnej miary Lebesgue'a.

Wynika to właśnie ze wspomnianego wcześniej Twierdzenia Ergodycznego oraz faktu, że T jest przekształceniem zachowującym miarę Lebesgue'a i ergodycznym[3]. O ile to pierwsze jest łatwym do wykonania ćwiczeniem dla Czytelnika, o tyle to drugie chyba znów nie ma elementarnego dowodu. W zamian podamy w końcu sformułowanie Twierdzenia Ergodycznego Birkhoffa.

Twierdzenie 1 (Ergodyczne Birkhoffa). *Niech $T : X \to X$ będzie przekształceniem zachowującym miarę μ i ergodycznym, zaś $\phi : X \to \mathbb{R}$ funkcją całkowalną. Wtedy:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(T^i(x)) = \int \phi \, d\mu$$

dla μ-prawie wszystkich x.

Warto zauważyć, że to twierdzenie jest dokładnym odpowiednikiem Mocnego Prawa Wielkich Liczb i można jest wyrazić w ten sposób, że ciąg zmiennych losowych $\phi_i = \phi \circ T^i$ spełnia MPWL. Twierdzenie Birkhoffa może zatem kojarzyć się z MPWL Kołmogorowa. Jest to częściowo słuszne, bo niezmienniczość układu jest równoważna z tym, że zmienne ϕ_i mają ten sam rozkład. Ergodyczność jednak nie jest „zamiennikiem" równoważności, bo na przykład istnieją przekształcenia ergodyczne, które nie spełniają Centralnego Twierdzenia Granicznego.

[3] ergodyczność znaczy np. to, że wszystkie funkcje T-niezmiennicze są stałe prawie wszędzie

Z drugiej strony istnieją układy „porządne”, dla których CTG i jeszcze mocniejsze własności statystyczne zachodzą. Bodaj najważniejszym dla teorii układów dynamicznych przykładem są tak zwane przesunięcia Bernoulliego, zdefiniowane na *przestrzeni symbolicznej*: $\Sigma := A^{\mathbb{N}}$ dla pewnego alfabetu A. Przesunięciem σ nazywamy przekształcenie:

$$\sigma((\omega_0, \omega_1, \omega_2, \ldots)) = (\omega_1, \omega_2, \ldots)$$

Istotność tego przykładu wynika z dwóch powodów.

Pierwszym jest to, że — wbrew pozorom — takie układy często występują „w przyrodzie”, na przykład nawinięcie T to nic innego niż przesunięcie Bernoulliego na przestrzeni słów nad alfabetem $\{0, 1\}$, o ile liczbę $\theta \in [0, 1)$ utożsamimy z ciągiem jej rozwinięcia binarnego.

Drugim powodem jest stosunkowa łatwość zrozumienia „jak działają” przesunięcia. Łatwo na przykład znaleźć ciąg $\omega \in \Sigma$ o dowolnym okresie. Równie łatwo wymyślić ciąg, którego trajektoria jest gęsta.[4]

Wracając do własności statystycznych: brakuje nam jeszcze miary niezmienniczej. Mając dowolną miarę m na alfabecie A możemy na Σ wziąć μ: nieskończony produkt tych miar, czyli po prostu nieskończony ciąg niezależnych „losowań” z rozkładem m. Gdyby zatem funkcja ϕ (por. Twierdzenie Ergodyczne) zależała tylko od pierwszego symbolu, to zmienne ϕ_i byłby po prostu niezależne, a zatem spełniałyby Centralne Twierdzenie Graniczne. Podobnie, gdyby zależały tylko od skończenie wielu symboli, to ϕ byłoby niezależne od ϕ_i dla odpowiednio dużych i i z tego również dałoby się „wycisnąć” CTG. Możemy próbować zatem przybliżać w ten sposób dowolne funkcje ϕ. Metody te mają, jak to w matematyce, wiele uogólnień, ale ich już rozważać nie będziemy i na tym skończymy ten krótki zarys związków probabilistyki z układami dynamicznymi…

[4] możemy przykładowo wypisać po kolei rozwinięcia binarne wszystkich liczb naturalnych