Teoria ergodyczna
WPPT Matematyka, IIIr. semestr zimowy 2008/9
Wykład 8

17/12/08

Układy symboliczne, rozbicia skończone, generatory, procesy Markowa

Układy symboliczne.
Niech $\Lambda = \{1, 2, \dots, l\}$. *Układem symbolicznym* nad alfabetem Λ nazywamy układ (Y, ν, S), gdzie $Y = \Lambda^{\mathbb{N}_0}$ (przypadek nieodwracalny) lub $Y = \Lambda^{\mathbb{Z}}$ (przypadek odwracalny), transformacja $S : Y \to Y$ to „szyft" zadany wzorem

$$(S(y))_n = y_{n+1},$$

gdzie $y = (y_n) \in Y$, oraz gdzie ν jest dowolną miarą probabilistyczną S-niezmienniczą (jest wiele takich miar, na przykład $\{p_1, \dots, p_l\}^{\mathbb{N}_0 \text{ (lub } \mathbb{Z})}$, gdzie $\{p_1, \dots, p_l\}$ jest wektorem probabilistycznym, a także wiele innych miar, które poznamy później).

Procesy generowane przez rozbicia skończone.
Niech (X, μ, T) będzie układem z miarą probabilistyczną i $T : X \to X$ zachowującym miarę. Niech $\mathcal{P} = \{A_1, A_2, \dots, A_l\}$ będzie mierzalnym rozbiciem skończonym (partycją) (tzn. zbiory A_i są rozłączne i ich suma ma miarę 1). Z rozbiciem tym związać można pewien *faktor symboliczny* (zwany też *procesem generowanym przez* $\mathcal{P}$) w następujący sposób: $Y = \Lambda^{\mathbb{N}_0}$, gdzie $\Lambda = \{1, 2, \dots, l\}$ (numery zbiorów rozbicia $\mathcal{P}$) (jeśli T jest odwracalne, to można też rozważać $\Lambda^{\mathbb{Z}}$), transformacja $S : Y \to Y$ to „szyft" i wreszcie odwzorowanie faktorujące $\pi : X \to Y$ zdefiniowane tak:

$$\pi(x) = y = (y_n), \quad \text{gdzie} \quad y_n = i \quad \iff \quad T^n(x) \in A_i.$$

(Taka sama definicja działa dla T odwracalnego i $n \in \mathbb{Z}$). Innymi słowy, punktowi x przypisujemy jego „historię wędrówki" po zbiorach A_i w postaci ciągu numerów tych zbiorów. Wreszcie na układ (topologiczny) (Y, S) przenosimy miarę μ odwzorowaniem π, $\nu = \pi(\mu)$ (czyli $\nu(B) = \mu(\pi^{-1}(B))$, B jest zbiorem mierzalnym w Y, czyli należy do sigma-ciała produktowego). Teraz sprawdza się elementarnie, że ν jest S-niezmiennicza, i że układ (Y, ν, S) jest faktorem układu (X, μ, T) przez odwzorowanie π (które jest ekwiwariantne).

Przykład 1 (dobrze nam znany). $(X, \mu, T) = (\mathbb{T}, \lambda, \cdot^2)$ (czyli zespolony okrąg jednostkowy z transformacją "podnoszenie do kwadratu"), $\mathcal{P} = \{(\overset{\frown}{-1, 1}], [\underset{\smile}{-1, 1})\}$ (czyli górny i dolny półokrąg). Wtedy $Y = \{0, 1\}^{\mathbb{N}_0}$ i $\nu = \{\frac{1}{2}, \frac{1}{2}\}^{\mathbb{N}_0}$. W tym przykładzie odzwzorowanie faktorujące π przyporządkowuje lczbie $e^{2\pi i t}$ rozwinięcie dwójkowe liczby t (nie jest to jednoznaczne tylko w punktach, które są pierwiastkami z jedności). Jest to odwzorowanie odwracalne prawie wszędzie, czyli układy (X, μ, T) i (Y, ν, S) są izomorficzne poprzez odwzorowanie π (oczywiście nie zawsze tak musi być).

Definicja. Powiemy, że $\mathcal{P}$ jest *generatorem* dla (X, μ, T) jeśli generowany proces jest izomorficzny z całym układem (X, μ, T) poprzez odwzorowanie π. Warunkiem równoważnym jest $\sigma(\bigcup_n \mathcal{P}^{-n}) = \mathcal{B}$ (gdzie $\mathcal{B}$ jest sigma-ciałem w X, równość jest modulo μ, a n przebiega $\mathbb{N}_0$ lub $\mathbb{Z}$, w zależności od tego jakie działanie rozważamy.)

Przykład 2. Weźmy układ symboliczny $(X, \mu, T) = (\Lambda^{\mathbb{N}_0}, \mu, S)$, gdzie μ jest miarą niezmienniczą na „szyft" S. Niech $\mathcal{P} = \{A_1, \dots, A_l\}$ będzie rozbiciem „na zerowej współrzędnej: $A_i = [i] = \{x = (x_n) : x_0 = i\}$. Wtedy proces indukowany przez $\mathcal{P}$ jest tożsamy z układem (X, μ, T). $\mathcal{P}$ jest oczywiście generatorem.

Uwagi.
1) Nie wystarczy aby proces generowany był izomorficzny z (X, μ, T) (poprzez jakieś inne odwzorowanie). Przykład: dowolny jednostronny układ symboliczny i rozbicie na pierwszej (a nie zerowej!) współrzędnej. Układ ten jest izomorficzny z całością (izomorfizmem jest odwzorowanie identycznościowe na ciągach), ale izomorfizmem nie jest naturalne odwzorowanie faktorujące, którym w tym wypadku jest shift S.
2) Układ odwracalny może mieć zarówno *generator obustronny* (tzn. takie rozbicie $\mathcal{P}$, że $\sigma(\bigcup_{n\in\mathbb{Z}} \mathcal{P}^{-n}) = \mathcal{B}$), jak i jednostronny (tzn. takie rozbicie $\mathcal{P}$, że $\sigma(\bigcup_{n\in\mathbb{N}_0} \mathcal{P}^{-n}) = \mathcal{B}$). Posiadanie generatora jednostronngo przez odwzorowanie odwracalne ma bardzo silne konsekwencje (a dwustronnego – dużo słabsze). Będzie o tym mowa przy okazji omawiania entropii.

Przykład 3. $(X, \mu, T) = ([0,1)^{\mathbb{Z}}, \lambda^{\mathbb{Z}}, S)$ (tutaj $S = S_I$ oznacza "szyft" na przestrzeni ciągów o wartościach w $[0,1)$), $\mathcal{P} = \{C_1, C_2, \dots, C_l\}$, $C_i = \{x = (x_n) : x_0 \in [\frac{i-1}{l}, \frac{i}{l})\}$. Teraz faktorem generowanym jest $(\{1, 2, \dots, l\}^{\mathbb{Z}}, \{\frac{1}{l}, \frac{1}{l}, \dots, \frac{1}{l}\}^{\mathbb{Z}}, S)$ (tu z kolei $S = S_l$ oznacza "szyft" na przetrzeni ciągów o wartościach w $\{1, 2, \dots, l\}$ – od tej pory wszelkie „szyfty" będziemy oznaczać po prostu literą S). Tym razem oczywiście $\mathcal{P}$ nie jest generatorem: nie rozróżniane są ciągi, które zawsze przechodzą przez te same „sektory".

Procesy Markowa.
Wprowadzimy teraz pewną ważną klasę miar niezmienniczych w układach symbolicznych. Są to *miary Markowskie.*
Niech $\Lambda = \{1, 2, \dots, l\}$. Niech M będzie macierzą stochastyczną $l \times l$ (to znaczy, każdy wiersz jest wektorem probabilistycznym). Wtedy:

Twierdzenie. Macierz M ta ma co najmniej jeden probabilistyczny lewostronny wektor niezmienniczy $p = (p_1, \dots, p_l)$ (tzn. taki, że $pM = p$). Jeśli macierz ta ma tę własność, że pewna potęga M^k ma w przynajmniej jednej kolumnie wszystkie współczynniki niezerowe, to wektor probabilistyczny niezmienniczy jest jedyny i dowolny wiersz macierzy M^n zmierza (po n) do tego wektora.

Dowód: Wektory probabilistyczne l-wymiarowe tworzą zbiór zwarty wypukły Δ. Mnożenie przez M (wektor z lewej, macierz z prawej) jest transformacją afiniczną i ciągłą Δ w siebie. Wiemy już, że takie odwzorowanie ma ZAWSZE co najmniej jeden punkt stały (korzystaliśmy z tego dowodząc istnienia miar niezmienniczych w układach topologicznych). Teraz załóżmy istnienie ściśle dodatniej kolumny dla pewnej potęgi M^k. Niech $p = (p_1, \dots, p_l)$ i $q = (q_1, \dots, q_l)$ będą dwoma różnymi wektorami probabilistycznymi. Wtedy $p - q$ jest wektorem o sumie zerowej. Niech $\|x\|$ oznacza sumę modułów współrzędnych wektora $x \in \mathbb{R}^l$. Policzmy

$$(*) \qquad \|xM^k\| = \sum_j \left| \sum_i x_i M^k_{i,j} \right| \le \sum_j \sum_i |x_i| M^k_{i,j} = \sum_i |x_i| \sum_j M^k_{i,j} = \sum_i |x_i| 1 = \|x\|.$$

Równość zachodzi wtedy i tylko wtedy gdy dla każdego j zachodzi równość

$$\left|\sum_i x_i M^k_{i,j}\right| \leq \sum_i |x_i| M^k_{i,j}.$$

Ale taka równość zachodzi tylko wtedy, gdy albo $M^k_{i,j} = 0$ dla wszystkich i, dla których x_i jest ujemne albo to samo dla wszystkich i dla których x_i jest dodatnie. Jeśli x jest wektorem niezerowym o sumie zero, istnieje przynajmniej jedno i dla którego $x_i < 0$ i przynajmniej jedno i takie, dla którego $x_i > 0$. Z założenia, istnieje j (numer kolumny) takie, że $M^k_{i,j} \neq 0$ dla wszystkich i, zatem w (*) wystąpi ostra nierówność. Powyższe rozumowanie pokazuje, że M^k zbliża do zera (w powyższej normie) wszystkie wektory o sumie zerowej. Łatwo widać, że współczynnik zmniejszania $c(x) = \frac{\|xM^k\|}{\|x\|}$ nie zależy od unormowania wektora, poza tym jest on funkcją ciągłą na sferze jednostkowej. Sfera ta jest zwarta, zatem funkcja $c(x)$ (ostro mniejsza od 1) osiąga swoje supremum c, które musi zatem być też mniejsze od 1. Czyli każdy niezerowy wektor o sumie zerowej zmniejsza normę co najmniej o współczynnik c. Wynika z tego, że mnożenie przez M jest odwzorowaniem zbliżającym na wektorach probabilistycznych (bo, przypomnijmy, $p - q$ jest wektorem o sumie zerowej). Z tw. Banacha, ma ono jedyny punkt stały p_0. Zauważmy, że wektor niezmienniczy dla M jest też niezmienniczy dla M^k. Ponieważ istnieje co najmniej jeden wektor M-niezmineniczy, musi to być p_0 i jest on jedyny. Wiemy też z Tw. Banacha, że p_0 jest granicą dowolnego ciągu pM^{kn}. W szczególności za p można wziąć dowolny wektor postaci pM^i i wtedy otrzymamy, że do p_0 zbiega pM^{i+kn}. Podstawiając kolejno $i = 1, \ldots, k-1$ wnioskujemy, że do p_0 dąży cały ciąg pM^n. Biorąc teraz za p j-ty wektor bazy standardowej ($j = 1, 2, \ldots, l$) otrzymamy zbieżność do p_0 j-tego wiersza macierzy M^n. $\square$

Uwaga: Szczególnym przypadkiem macierzy spełniającej założenie o ściśle dodatniej kolumne jest macierz *nieprzywiedlna*, tzn. taka, że M^k jest dla pewnego k ściśle dodatnia. W takim wypadku p_0 jest wektorem ściśle dodatnim.

Definicja. Miarą markowską na $\{1, \ldots, l\}^{\mathbb{N}_0}$ nazywamy miarę przyjmującą na cylindrach $[x_1, \ldots, x_n] = \{x = x_1, \ldots, x_n, \ldots\}$ wartości wg. wzoru

$$\mu([x_1, \ldots, x_n]) = (p_0)_{x_1} M_{x_1,x_2} M_{x_2,x_3} \cdots M_{x_{n-1},x_n},$$

gdzie M jest macierzą stochastyczną o wymiarach $l \times l$, a p_0 wektorem probabilistycznym lewostronnie niezmienniczym dla M. Takim samym wzorem zadaje się miarę na $\{1, \ldots, l\}^{\mathbb{Z}}$, z tą różnicą, że miarę zadaje się na cylindrach symetrycznych $[x_{-n}, \ldots, x_0, \ldots, x_n]$ i we wzorze wystąpi $(p_0)_{x_{-n}}$ (a potem iloczyn odpowiednich współczynników macierzy).

Twierdzenie. Miara μ jest poprawnie określona i niezmiennicza na ,,szyft".

Zadanie 1. Dowód przeprowadzić samodzielnie.

Interpretacja: Proces taki ilustruje spacer losowy po grafie o l wierzchołkach oznaczonych cyframi $1, 2, ..., l$. W chwili zerowej zaczynamy z wierzchołka wybranego losowo zgodnie z rozkładem p_0. Następnie w kolejnych chwilach z wierzchołka j, w którym się w danech chwili znajdujemy, wędrujemy do wierzchołka zgodnie z rozkładem zadanym przez j-ty wiersz macierzy M (czyli do wierzchołka k z prawdopodobieństwem $M_{j,k}$). Prawdopodobieństwo, że przejdziemy kolejno

$x_1, x_2, \ldots, x_n$ (a potem już dowolnie) jest dokładnie takie, jak mówi powyższy wzór na $\mu([x_1, \ldots, x_n])$.

Zadanie 2.
A) Wykaż, że jeśli dla pewnego $k \geq 1$ M^k posiada kolumne ściśle dodatnią, to dla każdych dwóch cylindrów A i B zachodzi $\mu(A \cap T^{-n}(B)) \to \mu(A)\mu(B)$ (gdzie μ jest miarą zadana przez jedyny wektor p_0).
B) Jak z tego wynika ergodyczność miary μ?

Tomasz Downarowicz