Wykład 8: Tw. ergodyczne. Łańcuchy odwracalne.

Wykładowca: Andrzej Ruciński *Pisarz:Przemysław Rogowski, Tomasz Rzędowski*

1 Wstęp

W pierwszej części dokończymy zagadnienia z poprzedniego wykładu ukazując zastosowania couplingu na kilku przykładach. Następnie przedstawione będzie twierdzenie ergodyczne. W drugiej części wprowadzimy odwracalne łańcuchy Markowa i podamy kilka prostych przykładów.

2 (dokończenie)

Twierdzenie 1. *(Cam 1960) Niech $X_1, X_2, ..., X_n$ będą niezależnymi zmiennymi losowymi o rozkładzie 0,1. Niech $\boldsymbol{E}X_r = p_r$ oraz $S = \sum_{r=1}^{n} X_r$. Wówczas*

$$d_{TV}(S, P) \leq \sum_{r=1}^{n} p_r^2$$

oraz P ma rozkład $Po(\lambda = \sum_{r=1}^{n} p_r)$

Dowód. Wprowadźmy coupling par niezależnych (X_r, Y_r) takich, że

$$P(X_r = x, Y_r = y) = \left\{ \begin{array}{ll} 1 - p_r & x = y = 0 \\ e^{-p_r} & x = 1,\ y = 0 \\ \frac{p_r^y}{y!} e^{-p_r} & x = 1,\ y \geq 1 \end{array} \right\}$$

Zauważmy, że X_r ma rozkład $Ber(p_r)$ oraz Y_r ma rozkład $Po(p_r)$. Postarajmy się ukazać ideę wprowadzenia takiego couplingu tym, którzy nie są wstanie dojrzeć jej już za pierwszym razem. Rozpatrzymy następującą równość

$$P(X_r \neq Y_r) = p_r - p_r e^{-p_r}.$$

W oszacowaniu jej skorzystamy z następującego lematu.

Lemat 1. *Dla każdego $x \in < 0, 1 >$, zachodzi*

$$e^{-x} \geq 1 - x.$$

Dowód. Niech $\varphi(x) = e^{-x} - 1 + x$. Każde dziecko widzi, że jest to funkcja różniczkowalna na przedziale $< 0, 1 >$. Otrzymujemy

$$\varphi(x') = -e^{-x} + 1 \leq 0,$$

gdyż $\forall_{x \in <0,1>} \quad e^{-x} \geq 1$. Jest to zatem funkcja malejąca na przedziale $< 0, 1 >$. Czyli $\varphi(x)$ przyjmuje najmniejszą wartość w dziedzinie w punkcie 1.

$$\forall_{x \in <0,1>} \quad \varphi(x) \geq \varphi(1) = e^{-1} \geq 0.$$

Czyli

$$e^{-x} - 1 + x \geq 0$$

$\square$

W oparciu o lemat otrzymujemy

$$\begin{aligned} P(X_r \neq Y_r) &= \\ = p_r - p_r e^{-p_r} &= \\ = p_r(1 - e^{-p_r}) \leq p_r p_r &= \\ &= p_r^2. \end{aligned}$$

Przypomnijmy, że

$$S = \sum_{n}^{r=1} X_r, \quad P = \sum_{r=1}^{n} Y_r$$

$$P \quad \text{ma rozkład} \quad Po(\lambda = \sum_{r=1}^{n} p_r)$$

$$d_{TV}(S, P) = \frac{1}{2} \sum_{k} |\mathbf{P}(S = k) - \mathbf{P}(P = k)|$$

Zauważmy, że

$$\begin{aligned} |\mathbf{P}(S = k) - \mathbf{P}(P = k)| = |\mathbf{P}(S = k, P \neq k) - \mathbf{P}(S \neq k, P = k)| \leq \\ \leq \mathbf{P}(S = k, P \neq k) + \mathbf{P}(S \neq k, P = k) \end{aligned}$$

Ostatnia nierówność wynika z faktu, że prawdopodobieństwo jest nieujemne.

Wróćmy na główny tor.

$$d_{TV}(S, P) \leq \frac{1}{2} 2\mathbf{P}(S \neq P) = \mathbf{P}(S \neq P)$$

Stąd otrzymujemy, w jakże oczywisty sposób, że

$$\mathbf{P}(S \neq P) \leq \mathbf{P}(\exists_r, X_r \neq Y_r) \leq \sum_{r=1}^{n} \mathbf{P}(X_r \neq Y_r) \leq \sum_{r=1}^{n} p_r^2$$

$\square$

W ten oto sposób zakończyliśmy podawanie przykładów zastosowań couplingu. Czas powrócić do potężnego Twierdzenia Ergodycznego. Przypomnijmy je.

Twierdzenie 2. *(Twierdzenie Ergodyczne) Dla każdego nieredukowalnego, nieokresowego łańcucha Markowa X oraz dla każdego rozkładu początkowego $\mu^{(0)}$ mamy $\mu^{(n)} \stackrel{TV}{\to} \pi$, gdzie π jest rozkładem stacjonarnym.*

Dowód. Udowodniliśmy już jedyność π na wcześniejszym wykładzie. Pozostaje wykazać zbieżność $\mu^{(n)} \stackrel{TV}{\to} \pi$. Niech $S = 1, 2, ..., k$ będzie zbiorem stanów. Musimy pokazać, że

$$\sum_{i=1}^{k} |\mu_i^{(n)} - \pi_i| \to 0$$

Niech $X = (X_n)$ będzie naszym łańcuchem Markowa. Zapuścimy drugi łańcuch $Y = (Y_n)$ mający te same stany ale inne warunki początkowe: X ma $\mu^{(0)}$, natomiast Y ma π.

Aby nie zagmatwać się w labiryncie trików przedstawmy główną ideę tego dowodu. Jesli $X_m = Y_m$ dla pewnego m, to X_n i Y_n mają ten sam rozkład od tego momentu. Wynika to stąd, że opisuje je ta sama macierz.

Wprowadźmy zmienną losową

$$T = min\{m : X_m = Y_m\}.$$

$$\forall_{i \in S} \forall_n \quad \mu_i^{(n)} = \mathbf{P}(X_n = i) = \mathbf{P}(X_n = i | T \leq n) + \mathbf{P}(X_n = i | T > n)$$

Zauważmy, że w wyrażeniu $\mathbf{P}(X_n = i | T \leq n)$ możemy wymienić X_n na Y_n, gdyż w tej przestrzeni warunkowej są takie same.

$$\begin{aligned} \mu_i^{(n)} \quad &= \mathbf{P}(Y_n = i | T \leq n) \mathbf{P}(T \leq n) + \\ &+ \mathbf{P}(X_n = i | T > n) \mathbf{P}(T > n) \leq \\ &\leq \mathbf{P}(Y_n = i) + \mathbf{P}(T > n) \end{aligned}$$

stąd

$$\mu_i^{(n)} \leq \pi_i + \mathbf{P}(T > n)$$

co implikuje

$$|\mu_i^{(n)} - \pi_i| \leq \mathbf{P}(T > n)$$

Pozostało jedynie pokazać, że $\mathbf{P}(T > n) \stackrel{n \to \infty}{\longrightarrow} 0$. □

Uwaga. Gdy

$$\mu^{(0)} = (0, \ldots, 0, 1, 0, \ldots, 0)$$

gdzie jedynka jest na i-tym miejscu, wtedy

$$\mu_i^{(0)} \to \pi_i$$

Innymi słowy, nie istotne jest skąd zaczynamy, ale dokąd zmierzamy. Po pewnej ilości kroków stan początkowy „rozpływa się” i od pewnego momentu nie ważne jest już skąd zaczynaliśmy.

3 Odwracalne Łańcuchy Markowa

W tej części wykładu wprowadzone zostaną odwracalne łańcuchy Markowa, proste własności i implikacje oraz przykłady.

Definicja 1. *Rozkład π nazywamy odwracalnym, gdy*

$$\forall_{i,j\in S} \quad \pi_i P_{ij} = \pi_j P_{ji}$$

gdzie π_i jest i-tą współrzędna wektora prawdopodobieństwa rozkładu π, a P_{ij} prawdopodobieństwem przejścia ze stanu s_i do stanu s_j.

Definicja 2. *Łańcuch Markowa nazywamy odwracalnym gdy jego rozkład jest dowracalny.*

Lemat 2. *Rozkład odwracalny π jest rozkładem stacjonarnym.*

Dowód.

$$\pi_j = \pi_j \cdot 1 = \pi_j \sum_{i=1}^{k} P_{ji} = \sum_{i=1}^{k} \pi_j P_{ji} = \sum_{i=1}^{k} \pi_i P_{ij}.$$

□

Lemat 3. *Niech X, Y będą zmiennymi losowymi takimi, że*

$$\forall_{n=0,1,\ldots,N} \quad Y_n = X_{N-n},$$

oraz niech $\mu^{(0)} = \pi$ będzie rozkładem odwracalnym. Wtedy

$$P(Y_{n+1} = j | Y_n = i) = \frac{\pi_j}{\pi_i} P_{ji} \quad (= P_{ij} = P(X_{n+1} = j | X_n = i))$$

Dowód. Pokażemy, że Y jest łańcuchem Markowa.

$$P(Y_{n+1} = i_{n+1} | Y_n = i_n, \ldots, Y_0 = i_0) =$$

ze wzoru na prawdopodobieństwo warunkowe mamy

$$= \frac{P(Y_j = i_j, 0 \leq j \leq n+1)}{P(Y_j = i, 0 \leq j \leq n)} =$$

podstawiając X_{N-n} za Y_n

$$= \frac{P(X_{N-n-1} = i_{n-1}, \ldots, X_N = i_0)}{P(X_{N-n} = i_n, \ldots, X_N = i_0)} =$$

stosując wzór łańcuchowy oraz $P_{i_n+1,i_n} = P_{i_n,i_n+1}$ otrzymujemy

$$\frac{\pi_{i_{n+1}} P_{i_n+1,i_n} \cdot \ldots \cdot P_{i_1,i_0}}{\pi_{i_n} P_{i_n,i_{n-1}} \cdot \ldots \cdot P_{i_1,i_0}} = \tag{1}$$

$$= \frac{\pi_{i_{n+1}} P_{i_n+1,i_n}}{\pi_{i_n}} = \tag{2}$$

$$= \frac{\pi_{i_n} P_{i_n,i_{n+1}}}{\pi_{i_n}} = \tag{3}$$

$$= P(N_{N-n-1} = i_{n+1} | X_{N-n} = i_n) \tag{4}$$

z założenia $X_{N-n-1} = Y_{n+1}$, stąd

$$P(Y_{n+1} = i_{n+1} | Y_n = i_n)$$

$\square$

Przykład 1. *Spacer po grafie G. Prawdopodobieństwo przejścia z wierzchołka i do wierzchołka j wynosi*

$$p_{ij} = \begin{cases} \frac{1}{d_i} & \text{gdy } ij \in E \text{ (tj. jest krawędzią, } d_i \text{ - stopień wierzchołka } i \text{)} \\ 0 & \text{gdy } ij \text{ nie jest krawędzią} \end{cases}$$

Jest to łańcuch Markowa odwracalny. Rozkład

$$\pi = (\frac{d_1}{2e}, \ldots, \frac{d_n}{2e})$$

jest rozkładem stacjonarnym i odwracalnym ponieważ

$$\pi_i P_{ij} = \frac{d_i}{2e} \frac{1}{d_i} = \frac{1}{2e} \tag{5}$$

$$\pi_j P_{ji} = \frac{d_j}{2e} \frac{1}{d_j} = \frac{1}{2e} \tag{6}$$

Przykład 2. *(proces narodzin i śmierci) Niech* $S = [k] = \{1, 2, 3, \ldots, k\}$

$$P_{ij} = \begin{cases} > 0 & \text{gdy } |i-j| = 1 \\ 0 & \text{w przeciwnym przypadku, tj. gdy } |i-j| \geq 2 \end{cases}$$

tj. prawdopodobieństwo przejścia ze stanu i do stanu j wynosi zero gdy stany te spełniają równość $|i-j| \geq 2$ *, w przeciwnym przypadku prawdopodobieństwo to jest dodatnie. Dla* $j = i$ *prawdopodobieństwo* P_{ii} *jest dowolne.*

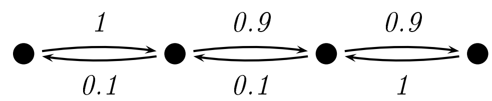
Otrzymujemy macierz przejścia

$$\begin{pmatrix} ? & >0 & & & \\ >0 & ? & >0 & & \\ & \ddots & \ddots & \ddots & \\ & & & & >0 \\ & & & >0 & ? \end{pmatrix}$$

Oto konkretny przykład dla $k = 4$. Macierz prawdopodobieństwa jest postaci

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 \\ 0 & 0.1 & 0 & 0.9 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

co daje graf przejść



Niech teraz π_1^* będzie dowolne, i niech

$$\pi_i^* = \frac{P_{i-1,i}}{P_{i,i-1}} \pi_{i-1}^*.$$

Rozkład π dany wzorem

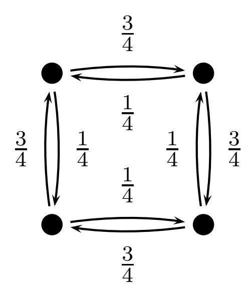
$$\pi_i = \frac{\pi_i^*}{\sum \pi_i^*}$$

jest rozkładem stacjonarnym. Łatwo sprawdzić, że w naszym przykładzie rozkład przyjmuje następującą postać :

$$\pi = \left(\frac{1}{182}, \frac{10}{182}, \frac{90}{182}, \frac{8}{182} \right)$$

Przykład 3. *Przykład łańcucha z rozkładem stacjonarnym i nieodwracalnym.*

Niech dany będzie łancuch Markowa



Rozkład $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ jest stacjonarny. Nie jest on odwracalny, ponieważ ma „tendencję” do „przemieszczania” się zgodnie z kierunkiem wskazuwek zegara. Brak symetryczności.