**RAP 412** 19.11.2008

# Wykład 7: Ciąg dalszy rozkładów stacjonarnych

*Wykładowca: Andrzej Ruciński* *Pisarze:Arkadiusz Buchelt, Przemysław Sokołowski*

## Wstęp

W tym wykładzie pokażemy istnienie rozkładu stacjonarnego dla dowolnego nierozkładalnego i nieokresowego łańcucha Markowa oraz jego postać. Korzystając ze zbieżności wykażemy jedyność owego rozkładu. Przedstawimy również metodę Coupling-u, która umożliwi wykazanie zbieżności do rozkładu stacjonarnego (stabilizowanie się łańcucha).

Przypomnijmy, że przestrzeń stanów $S$ jest skończona, to znaczy $S = \{s_1, s_2, \ldots, s_k\}$, oraz $T_{i,j} = \min\{n \geq 1 \colon X_n = s_j \wedge X_0 = s_i\}$ i $\tau_{i,j} = \mathbf{E}(T_{i,j} \mid X_0 = s_i)$.

**Lemat 1.** *Dla dowolnego nieredukowalnego i nieokresowego łańcucha Markowa, dla każdych stanów $s_i, s_j$ mamy, że*

$$\mathbf{P}(T_{i,j} < \infty \mid X_0 = s_i) = 1$$

*oraz*

$$\tau_{i,j} = \mathbf{E}(T_{i,j} \mid X_0 = s_i) < \infty.$$

*Dowód.* Wiemy, że istnieje $M$ takie, że macierz przejść $P^M$ składa się z dodatnich wartości, to znaczy istnieje takie $M$, że począwszy od momentu $M$ z dodatnim prawdopodobieństwem osiągniemy ze stanu $s_i$ stan $s_j$. Oznaczmy przez $\alpha = \min_{i,j} \left\{ P^M_{i,j} \right\} > 0$, wówczas otrzymujemy, że prawdopodobieństwo nie osiągnięcia stanu $s_j$ w mniej niż $M$ krokach wynosi

$$\mathbf{P}(T_{i,j} > M \mid X_0 = s_i) \leq \mathbf{P}(X_M \neq s_j \mid X_0 = s_i) \leq 1 - \alpha.$$

W dalszym ciągu pomocnym będzie poniższe szacowanie prawdopodobieństwa warunkowego, że stan $s_j$ nie zostanie osiągnięty w mniej niż $2M$ krokach, gdy w mniej niż $M$ krokach nie wystąpił ten stan

$$\mathbf{P}(T_{i,j} > 2M \mid T_{i,j} > M, X_0 = s_i) \leq \mathbf{P}(X_{2M} \neq s_j \mid T_{i,j} > M, X_0 = s_i) =$$
$$= \sum_{\substack{l=1 \\ l \neq j}}^{k} \mathbf{P}(X_{2M} \neq s_j \mid T_{i,j} > M, X_M = s_l, X_0 = s_i) = \sum_{\substack{l=1 \\ l \neq j}}^{k} \mathbf{P}(X_{2M} \neq s_j \mid T_{i,j} > M, X_M = s_l) =$$
$$= \sum_{\substack{l=1 \\ l \neq j}}^{k} \mathbf{P}(X_{2M} \neq s_j \mid X_M = s_l)\, \mathbf{P}(X_M = s_l \mid T_{i,j} > M) \leq (1 - \alpha) \sum_{\substack{l=1 \\ l \neq j}}^{n} \mathbf{P}(X_M = s_l \mid T_{i,j} > M) \leq$$
$$\leq 1 - \alpha.$$

Oszacujmy teraz prawdopodobieństwo, że stan $s_j$ nie zostanie osiągnięty w mniej niż $2M$ krokach

$$\begin{aligned}&\mathbf{P}\left(T_{i,j} > 2M \mid X_0 = s_i\right) \overset{\star}{=} \mathbf{P}\left(T_{i,j} > M \mid X_0 = s_i\right) \mathbf{P}\left(T_{i,j} > 2M \mid T_{i,j} > M, X_0 = s_i\right) \leq \\ &\leq (1-\alpha)(1-\alpha) = (1-\alpha)^2\end{aligned}$$

$\star$ ze wzoru łańcuchowego.

Powtarzając analogiczne rozumowanie dla prawdopodobieństw $\mathbf{P}\left(T_{i,j} > 3M \mid T_{i,j} > M, X_0 = s_i\right)$, $\mathbf{P}\left(T_{i,j} > 4M \mid T_{i,j} > M, X_0 = s_i\right)$ itd. otrzymujemy, że dla dowolnego $l \in \mathbb{N}$

$$\mathbf{P}\left(T_{i,j} > lM \mid X_0 = s_j\right) \leq (1-\alpha)^l,$$

zatem prawdopodbieństwo, że stan $s_j$ nie zostanie nigdy osiągnięty wynosi

$$\begin{aligned}&\mathbf{P}\left(T_{i,j} = \infty \mid X_0 = s_j\right) = \mathbf{P}\left(\bigcap_{l=1}^{\infty} \{T_{i,j} > lM \mid X_0 = s_j\}\right) = \lim_{l\to\infty} \mathbf{P}\left(T_{i,j} > lM \mid X_0 = s_j\right) \leq \\ &\leq \lim_{l\to\infty} (1-\alpha)^l = 0.\end{aligned}$$

Na zakończenie obliczmy oczekiwaną ilość kroków do osiągnięcia stanu $s_j$

$$\begin{aligned}\tau_{i,j} &\overset{\star\star}{=} \sum_{n=1}^{\infty} \mathbf{P}\left(T_{i,j} > n\right) = \sum_{l=0}^{\infty} \sum_{n=lM}^{(l+1)M-1} \mathbf{P}\left(T_{i,j} > n\right) \leq \sum_{l=0}^{\infty} \sum_{n=lM}^{(l+1)M-1} \mathbf{P}\left(T_{i,j} > lM\right) = \\ &= M \sum_{l=0}^{\infty} \mathbf{P}\left(T_{i,j} > lM\right) \leq M \sum_{l=0}^{\infty} (1-\alpha)^l = \frac{M}{\alpha} = const.\end{aligned}$$

Równość $\star\star$ wynika z twierdzenia 11. z rozdziału §5.6 z Wstęp do teorii prawdopodobieństwa J.Jakubowski, R.Sztencel. $\square$

Oznaczmy przez $\mu_j = \tau_{jj}$ czas oczekiwania na pierwszy powrót do stanu $s_j$ gdy zaczynamy z tego stanu.

**Twierdzenie 1.** *Dla dowolnego nieredukowalnego i nieokresowego łańcucha Markowa istnieje rozkład stacjonarny* $\mathbf{\Pi}$.

*Dowód.* Dla uproszczenia wprowadźmy następujące oznaczenia:

$$\begin{aligned}i &:= s_i,\\ T_l &:= T_{l,l}.\end{aligned}$$

Ustalmy dowolny stan $l \in [k]$. W dalszej części wszystkie zdarzenia są rozpatrywane w przestrzeni warunkowej z warunkiem $X_0 = s_l$. Dla każdego $i \in [k]$ będącego pewnym stanem definiujemy zmienną losową

$$N_i(l) = \sum_{n=1}^{\infty} I_{\{X_n = i\} \cap \{T_l \geq n\}}.$$

Jeśli $i \neq l$ to powyższa zmienna losowa liczymy ile razy odwiedzony zostanie stan $i$ po drodze do $l$. Oznaczmy przez $\rho_i(l) = \mathbf{E}(N_i(l) \mid X_0 = l)$

Uwaga oczywiście $N_l(l) = 1$, czyli również $\rho_i(i) = 1$. Zauważmy, że:

$$T_l = \sum_{i=1}^{k} N_i(l),$$

gdyż wszystkie kroki jakie pokonamy idąc z $l$ do $l$ możemy rozbić ze względu na liczbę odwiedzin w każdym ze stanów. Zatem

$$\mathbf{E}T_l = \mu_l = \sum_{i=1}^{k} \rho_i(l),$$

i na mocy lematu 1 $\mu_l < \infty$, czyli dla każdego $i \in [k]$ mamy $\rho_i(l) < \infty$.

Kandydatem na rozkład dla każdego $l$ jest

$$\mathbf{\Pi}(l) = \frac{\boldsymbol{\rho}(l)}{\mu_l},$$

gdzie $\boldsymbol{\rho}(l) = (\rho_1(l), \ldots, \rho_k(l))$, czyli $\mathbf{\Pi}(l) = (\Pi_1, \Pi_2, \ldots, \Pi_k) = \left(\frac{\rho_1(l)}{\mu_l}, \ldots, \frac{\rho_k(l)}{\mu_l}\right)$. Pokażemy, że jest to rozkład stacjonarny.

Mamy $N_i(l) \geq 0$, więc $\rho_i(l) \geq 0$, a stąd $\Pi_i(l) \geq 0$ oraz

$$\sum_{i=1}^{k} \Pi_i(l) = \sum_{i=1}^{k} \frac{\rho_i(l)}{\mu_l} = \frac{1}{\mu_l} \sum_{i=1}^{k} \rho_i(l) = 1$$

Zatem $\mathbf{\Pi}(l)$ jest rozkładem prawdopodobieństwa. Ponadto

$$\begin{aligned}
\rho_i(l) &= \sum_{n=1}^{\infty} \mathbf{P}(X_n = i, T_l \geq n \mid X_0 = l) = p_{li} + \sum_{n \geq 2} \sum_{\substack{j \neq 1 \\ j=1}}^{k} \mathbf{P}(X_n = i, T_l \geq n, X_{n-1} = j \mid X_0 = l) = \\
&= \rho_{li} + \sum_{n \geq 2} \sum_{j \neq 1} \mathbf{P}(X_{n-1} = j, T_l \geq n) \mathbf{P}(X_n = i \mid T_l \geq n, X_{n-1} = j) = \\
&= p_{li} + \sum_{n \geq 2} \sum_{j \neq l} \mathbf{P}(X_{n-1} = j, T_l \geq n-1) \mathbf{P}(X_n = i \mid X_{n-1} = j) = \\
&= \rho_l(l) p_{li} + \sum_{\substack{j \neq l \\ j=1}}^{k} p_{ji} \sum_{n=2}^{\infty} \mathbf{P}(X_{n-1} = j, T_l \geq n) = \rho_l(l) p_{li} + \sum_{\substack{j=1 \\ j \neq l}}^{k} p_{ji} \rho_j(l) \\
&= \sum_{j=1}^{k} p_{ji} \rho_j(l),
\end{aligned}$$

stąd

$$\boldsymbol{\rho}(l) = \boldsymbol{\rho}(l) P \ / \cdot \frac{1}{\mu_l}$$
$$\mathbf{\Pi}(l) = \mathbf{\Pi}(l) P$$

Co dowodzi, że $\mathbf{\Pi}(l)$ jest rozkładem stacjonarnym. $\square$

### Postać rozkładu stacjonarnego

W dalszej części pokażemy, że zachodzi następujący ciąg implikacji: zbieżność $\Rightarrow$ jedyność $\Rightarrow \frac{\rho_i(l)}{\mu_l} = \frac{\rho_i(i)}{\mu_i} = \frac{1}{\mu_i} \Rightarrow \mathbf{\Pi} = \left(\frac{1}{\mu_1}, \frac{1}{\mu_2}, \dots, \frac{1}{\mu_k}\right)$ jest rozkładem stacjonarnym.

**Twierdzenie 2.** *(Twierdzenie ergodyczne o zbieżności) Jeśli* $\mathbf{\Pi}$ *jest rozkładem stacjonarnym, to dla każdego* $\mu^{(0)}$ *mamy*

$$\mu^{(n)} \stackrel{TV}{\to} \mathbf{\Pi} \Leftrightarrow \lim_{n\to\infty} d_{TV}\left(\mu^{(0)}, \mu^{(n)}\right) = 0$$

*Dowód.* Dowód na kolejnym wykładzie. □

Na potrzeby dalszej części tego rozdziału załóżmy prawdziwość powyższego twierdzenia i pokażemy jedyność rozkładu stacjonarnego. Przypuśćmy zatem, że istnieją dwa rozkłady stacjonarne $\mathbf{\Pi}$ i $\mathbf{\Pi}'$. Niech $\mu^{(0)} = \mathbf{\Pi'}$, wtedy $\mu^{(n)} = \mathbf{\Pi}'$ dla każdego $n \in \mathbb{N}$, ponieważ jest on rozkładem stacjonarnym. Ze zbieżności (którą jak na razie tylko zakładamy)

$$\mu^{(0)} \stackrel{TV}{\to} \mathbf{\Pi} \Leftrightarrow \lim_{n\to\infty} d_{TV}\left(\mathbf{\Pi}', \mathbf{\Pi}\right) = 0,$$

a zatem $d_{TV}\left(\mathbf{\Pi}', \mathbf{\Pi}\right) = 0 \Rightarrow \mathbf{\Pi} = \mathbf{\Pi}'$. W celu udowodnienia zbieżności rozkładów wprowadźmy pojęcie couplingu.

## Coupling

Załóżmy, że dane są dwie zmienne losowe $X$, $Y$ na przestrzeniach probabilistycznych odpowiednio $\Omega_1$, $\Omega_2$. Konstruujemy „kopie” $X', Y'$ zmiennych $X$ i $Y$ na tej samej przestrzeni $\Omega$ tak, aby ich rozkłady brzegowe były takie same jak rozkłady $X$ i $Y$ oraz powinny one być ze sobą powiązane, dlatego między innymi przestrzenie produktowe w tym przypadku są praktycznie bezużyteczne.

**Przykład 1.** *(Stochastyczna dominacja)* $X$ *dominuje* $Y$*, gdy dla każdego* $x$ *wartości dystrybuant spełniają nierówność* $F_X(x) \leq F_Y(x)$*, czyli* $\mathbf{P}(X \leq x) \leq \mathbf{P}(Y \leq x)$ *co implikuje, że* $X \geq Y$ *dla dowolnej przestrzeni* $\Omega$*. W dalszej części będziemy wykorzystywać oznaczenie* $X \geq_{st} Y$.

**Twierdzenie 3.** *Jeśli* $X \geq_{st} Y$*, to istnieje* $(\Omega, \mathcal{F}, P)$ *i istnieją* $X'$ *i* $Y'$ *na* $(\Omega, \mathcal{F}, P)$ *takie, że* $F_{X'} = F_X$*,* $F_{Y'} = F_Y$ *oraz* $\mathbf{P}(X' \geq Y') = 1$.

*Piękny dowód.* Niech $\Omega = \langle 0, 1\rangle$, $\mathcal{F} = \mathcal{B}_{\mathbb{R}}(\langle 0, 1\rangle)$ oraz $P = \lambda_1(\langle 0, 1\rangle)$. Dla dowolnej dystrybuanty $F$ zdefiniujemy zmienną losową $Z_F$ na $\Omega$ w następujący sposób

$$Z_F(\omega) = \inf\{z \colon \omega \leq F(z)\}.$$

Zauważmy, że $\omega \leq F(z) \Leftrightarrow Z_F(\omega) \leq z$ stąd

$$F_{Z_F}(z) = \mathbf{P}(Z_F \leq z) = \mathbf{P}(\langle 0, F(z)\rangle) = \lambda_1(\langle 0, F(z)\rangle) = F(z).$$

Niech $X' = Z_{F_X}$, $Y' = Z_{F_Y}$ a wtedy $X' \geq Y'$ prawie wszędzie, czyli $\mathbf{P}(X' \geq Y') = 1$, co kończy dowód. □

**Przykład 2.** *(Zbieżność poissonowska) Wiemy, że* $Bin\left(n, \frac{\lambda}{n}\right) \to Po\left(\lambda\right)$*, gdy* $n \to \infty$ *(zbieżność szeregów liczbowych dla* $k \in \{0, \ldots, n\}$*). Jak dobre jest to przybliżenie?*

**Twierdzenie 4.** *(Cam, 1960) Jeśli* $X_1, \ldots, X_n$ *są niezależnymi zmiennymi losowymi Bernoulliego, gdzie* $\mathbf{E}X_i = p_i$*. Oznaczmy*

$$S = \sum_{r=1}^{n} X_r.$$

*Wtedy*

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

*gdzie*

$$Y \sim Po\left(\lambda = \sum_{r=1}^{n} p_r\right).$$

*W szczególności jeżeli* $p_r = \frac{\lambda}{n}$*, to* $d_{TV}\left(S, Y\right) \leq \frac{\lambda^2}{n}$*.*