

Spektrum i diagonalizacja łańcuchów Markowa

Wojciech Pachowiak

Wydział Matematyki i Informatyki
Uniwersytet Adama Mickiewicza w Poznaniu

Część 1: Łańcuchy Markowa

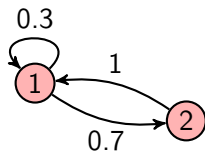
Część 2: Powtórka z algebry liniowej

Część 3: Twierdzenie Perrona-Frobeniusa

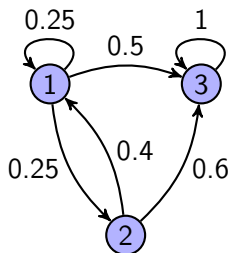
Część 4: Diagonalizacja odwracalnych łańcuchów Markowa

Przykłady łańcuchów Markowa

$$P = \begin{bmatrix} 0.3 & 0.7 \\ 1 & 0 \end{bmatrix}$$



$$P = \begin{bmatrix} 0.25 & 0.25 & 0.5 \\ 0.4 & 0 & 0.6 \\ 0 & 0 & 1 \end{bmatrix}$$



Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie przestrzenią probabilistyczną. Czym są σ -algebra $\mathcal{F}$ i miara probabilistyczna $\mathbb{P}$ będzie jasne z kontekstu.

Skończony zbiór $\mathcal{X} = \{1, 2, \dots, n\}$ nazywamy **przestrzenią stanów**.

Łańcuch Markowa możemy przedstawić jako:

- ▶ Ciąg zmiennych losowych $(X_0, X_1, X_2, \dots)$, gdzie $X_i : \Omega \rightarrow \mathcal{X}$.
- ▶ Skierowany graf z wagami $G = (\mathcal{X}, E, w)$, którego krawędzie wychodzące sumują się do 1.
- ▶ Macierz przejścia $P : \mathcal{X} \rightarrow \mathcal{X}$, która jest **kwadratowa**, **nieujemna** ($\forall x, y \in \mathcal{X} : P(x, y) \geq 0$) oraz **stochastyczna** ($\forall x \in \mathcal{X} \sum_{y \in \mathcal{X}} P(x, y) = 1$).

Intuicyjnie, zdarzenie elementarne $\omega \in \Omega$ jest ścieżką złożoną z wierzchołków grafu G , czyli $(X_0(\omega), X_1(\omega), X_2(\omega), \dots)$.

Interesują nas łańcuchy **jednorodne**, w których prawdopodobieństwa przejścia nie zmieniają się w czasie: $\forall a, b \in \mathcal{X}, \forall t \in \mathbb{N}_0$ mamy

$$\mathbb{P}(X_{t+1} = a \mid X_t = b) = \mathbb{P}(X_1 = a \mid X_0 = b).$$

Każdy łańcuch spełnia **własność Markowa**, która mówi, że przyszłość zależy tylko od stanu teraźniejszego: $\forall t \in \mathbb{N}_0, \forall x_i \in \mathcal{X}$ mamy

$$\begin{aligned} \mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_0 = x_0) \\ = \mathbb{P}(X_{t+1} = x_{t+1} \mid X_t = x_t) \end{aligned}$$

Fakt. Niech P będzie macierzą przejścia. Wtedy dla każdej pary stanów $x, y \in \mathcal{X}$ i każdego czasu $t \geq 0$ mamy

$$P(x, y) = \mathbb{P}(X_1 = y \mid X_0 = x)$$

$$P^t(x, y) = \mathbb{P}(X_t = y \mid X_0 = x).$$

Innymi słowy, $P^t(x, y)$ mówi nam o prawdopodobieństwie przejścia z x do y w t krokach.

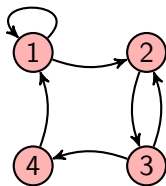
Oczywiście $P^0(x, x) = 1$ oraz $P^0(x, y) = 0$.

Definicja. Macierz przejścia P jest **nierozkładalna** wtw dla każdej pary stanów $x, y \in \mathcal{X}$ istnieje czas $t \geq 0$ taki, że $P^t(x, y) > 0$.

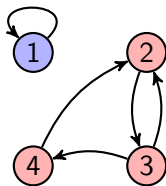
Intuicja. Z każdego stanu możemy dostać się do dowolnego innego stanu w skończonej liczbie kroków.

Obserwacja. W języku teorii grafów nierozkładalność oznacza, że podległy graf macierzy P jest *silnie spójny*.

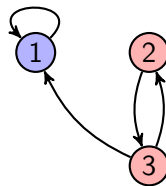
Nierozkładalność [PRZYKŁADY]



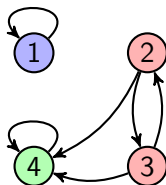
(a) Nierozkładalny



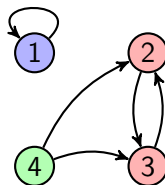
(b) Rozkładalny; dwie nierozkładalne składowe spójne



(c) Rozkładalny



(d) Rozkładalny; jedna składowa nierozkładalna, druga rozkładalna



(e) Tak samo jak w (c)

Definicja. Liczbę $d(x) := \gcd\{t \geq 1 : P^t(x, x) > 0\}$ nazywamy **okresem stanu** x .

Fakt. Niech P będzie nierozkładalna. Jeżeli jakiś stan ma okres $k \in \mathbb{N}$, to wszystkie stany mają okres k .

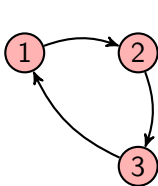
Wniosek. Okres jest własnością całego nierozkładalnego łańcucha Markowa.

Definicja. Jeżeli okres wszystkich stanów wynosi 1, to łańcuch nazywamy **nieokresowym**.

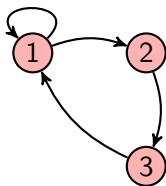
Obserwacja. Okres P możemy równoważnie zdefiniować jako liczbę:

$$\gcd\{k \in \mathbb{N} : \text{podległy graf } P \text{ zawiera skierowany cykl długości } k\}$$

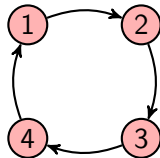
Okres i okresowość [PRZYKŁADY]



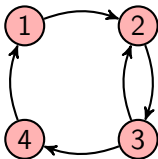
(a) Okres 3



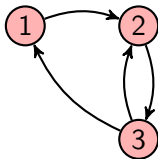
(b) Okres 1



(c) Okres 4



(d) Okres 2



(e) Okres 1



(f) Okres 1

Rozkład stacjonarny π

Definicja. Niech $\pi \in [0, 1]^{|X|}$ będzie wektorem wierszowym reprezentującym rozkład prawdopodobieństwa ($\sum_{x \in X} \pi(x) = 1$). Jeżeli dla każdego $y \in X$ spełniona jest równość

$$\pi(y) = \sum_{x \in X} \pi(x)P(x, y),$$

to π nazywamy *rozkładem stacjonarnym* łańcucha P .

Obserwacja. Rozkład stacjonarny π jest lewym wektorem własnym macierzy przejścia P odpowiadającym wartości własnej równej 1:

$$\pi P = \pi.$$

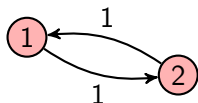
Fakt. Jeśli π jest rozkładem stacjonarnym P , to dla każdego $t \geq 0$ zachodzi również

$$\pi P^t = \pi.$$

Rozkład stacjonarny π [PRZYKŁADY]

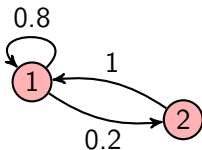
$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\pi = [1/2 \quad 1/2]$$



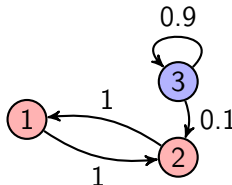
$$P = \begin{bmatrix} 0.8 & 0.2 \\ 1 & 0 \end{bmatrix}$$

$$\pi = [5/6 \quad 1/6]$$



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

$$\pi = [1/2 \quad 1/2 \quad 0]$$



Ważne twierdzenia!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

Twierdzenie. Jeśli P jest nierozkładalna, to istnieje **unikalny i ściśle dodatni** rozkład stacjonarny π .

(Zatem jest wtedy tylko jedno rozwiązanie równania $\pi P = \pi$)

Twierdzenie. Jeśli P jest nieredukowalna i nieokresowa, to

$$\lim_{t \rightarrow \infty} P^t(x, y) \rightarrow \pi(y).$$

(Wiersze P upodabniają się do π)

Fakt/Wniosek. Jeśli P jest nieredukowalna i nieokresowa, to $\exists N \in \mathbb{N}, \forall x, y \in \mathcal{X}, \forall n > N$ mamy $P^n(x, y) > 0$.

(Od pewnego momentu macierz staje się **dodatnia** i już taka pozostaje)

Jeżeli dla każdej pary $x, y \in \mathcal{X}$ rozkład prawdopodobieństwa $\nu \in [0, 1]^{|\mathcal{X}|}$ spełnia

$$\nu(x)P(x, y) = \nu(y)P(y, x), \quad (*)$$

to macierz przejścia P jest **odwracalna** względem ν .

Fakt. Jeśli P jest odwracalna względem ν , to ν jest rozkładem stacjonarnym P (czyli $\nu = \pi$).

Dowód. Sumujemy obustronnie (*):

$$\sum_{x \in \mathcal{X}} \nu(x)P(x, y) = \sum_{x \in \mathcal{X}} \nu(y)P(y, x) = \nu(y) \left[\sum_{x \in \mathcal{X}} P(y, x) \right] = \nu(y). \text{ Czyli } \nu P = \nu. \quad \square$$

Fakt. Jeśli P jest odwracalna, to jej okres jest ≤ 2 .

Dowód. Warunek (*) wymusza istnienie cyklu długości 2. □

Część 1: Łańcuchy Markowa

Część 2: Powtórka z algebry liniowej

Część 3: Twierdzenie Perrona-Frobeniusa

Część 4: Diagonalizacja odwracalnych łańcuchów Markowa

Niech A będzie macierzą kwadratową o wymiarze n , zbiorze wartości własnych Λ_A .

Definicja. Niech $\lambda_1, \lambda_2, \dots, \lambda_n \in \Lambda_A$ będą wartościami własnymi A , natomiast $f_1, f_2, \dots, f_n$ ich prawymi wektorami własnymi. Mówimy, że A jest **diagonalizowalna**, jeżeli zachodzi równość

$$A = NDN^{-1},$$

gdzie $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ jest macierzą diagonalną złożoną z wartości własnych, natomiast $N = [f_1 \ f_2 \ \dots \ f_n]$ jest odwracalną macierzą, zawierającą w swoich kolumnach prawe wektory własne A .

Fakt. Jeżeli wektory własne $f_1, f_2, \dots, f_n$ są liniowo niezależne, to A jest diagonalizowalna.

Wniosek. Jeżeli dla każdej wartości własnej jej wielokrotność algebraiczna i geometryczna są równe, to A jest diagonalizowalna.

Fakt. Jeżeli wartości własne A są różne (czyli $|\Lambda_A| = n$), to A jest diagonalizowalna.

Obserwacja. Ponieważ $N^{-1}A = DN^{-1}$, więc N^{-1} zawiera w swoich wierszach lewe wektory własne A . Oznaczmy je przez $u_1, u_2, \dots, u_n$. Mamy $N^{-1} = [u_1 \ u_2 \ \dots \ u_n]^T$.

Obserwacja. Postać lewych wektorów własnych zależy od wyboru prawych wektorów własnych (ponieważ N^{-1} zależy od N).

Obserwacja. Ponieważ $N^{-1}N = I$, więc

$$u_i^T f_j = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

Obserwacja. Równanie $A = NDN^{-1}$ można równoważnie zapisać jako

$$A(x, y) = \sum_{i=1}^n f_i(x) u_i(y) \lambda_i.$$

Diagonalizacja macierzy symetrycznych

Twierdzenie spektralne dla macierzy symetrycznych. Jeżeli M jest rzeczywistą macierzą symetryczną ($M = M^T$), to wartości własne M są rzeczywiste (czyli $\Lambda_M \subset \mathbb{R}$) i M jest diagonalizowalna w następujący sposób:

$$M = NDN^T,$$

gdzie N jest macierzą ortogonalną (czyli $NN^T = I$).

Fakt. Prawe i lewe wektory własne są rzeczywiste.

Obserwacja. Ponieważ $N^T = N^{-1}$, więc prawe i lewe wektory własne M są sobie równe (czyli $f_i = u_i$ dla każdego $i \in \{1, 2, \dots, n\}$).

Obserwacja. Ponieważ $N^T N = I$, więc

$$f_i^T f_j = \begin{cases} 0 & i \neq j \\ 1 & i = j, \end{cases}$$

czyli wektory własne tworzą bazą ortonormalną.

Sprzężone wartości własne

Fakt. Jeżeli A jest rzeczywista i $\lambda \in \Lambda_A$, to sprzężenie $\bar{\lambda}$ też jest wartością własną A .

Dowód. Zauważmy, że

$$\overline{Av} = \overline{\sum_j A(\cdot, j)v(j)} = \sum_j \overline{A(\cdot, j)v(j)} = \sum_j A(\cdot, j)\overline{v(j)} = A\bar{v}.$$

Zatem

$$Av = \lambda v$$

$$\overline{Av} = \overline{\lambda v}$$

$$A\bar{v} = \bar{\lambda}\bar{v}.$$



Część 1: Łańcuchy Markowa

Część 2: Powtórka z algebry liniowej

Część 3: Twierdzenie Perrona-Frobeniusa

Część 4: Diagonalizacja odwracalnych łańcuchów Markowa

Twierdzenie Perrona-Frobeniusa, cz. 1

Niech P będzie nieredukowalną macierzą przejścia o zbiorze wartości własnych Λ_P o postaci $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{|\mathcal{X}|}|$. Wtedy:

- ▶ $\lambda_1 = 1$
- ▶ Dla każdego $\lambda \in \Lambda_P \setminus \{\lambda_1\}$ zachodzi nierówność $\lambda_1 \geq |\lambda|$. Jeśli P jest nieokresowa to nierówność jest ostra: $\lambda_1 > |\lambda|$.
- ▶ Wielokrotność algebraiczna i geometryczna λ_1 wynosi 1.

Wniosek. Rozkład stacjonarny π jest unikalny.

- ▶ Jeżeli okres P wynosi $d \in \mathbb{N}$ i $\lambda \in \Lambda_P$, to $\lambda e^{2\pi i/d} \in \Lambda_P$. Czyli zbiór Λ_P jest niezmienniczy ze względu na obrót płaszczyzny zespolonej o kąt $2\pi/d$.
- ▶ Jeśli macierz A powstała przez odjęcie $0 < \epsilon < 1$ od dowolnego współczynnika macierzy P , to wartości własne $\lambda \in \Lambda_A$ mają moduł mniejszy od 1 (czyli $|\lambda| < 1$). Taką macierz nazywamy *substochastyczną*.

Niech A będzie nierozkładalną macierzą przejścia. Co możemy powiedzieć o wielokrotności algebraicznej wartości własnej $\lambda = 0$, jeżeli A ma...

Zadanie 1. ...rozmiar 5×5 i okres równy 3?

ODPOWIEDŹ: Jej wielokrotność jest równa 2.

Zadanie 2. ...rozmiar 547×547 i okres równy 9?

ODPOWIEDŹ: Jej wielokrotność jest większa lub równa 7.

Zadanie 3. ...rozmiar 6×6 , okres równy 2 i wiemy, że $0.22 + 0.1i$ jest wartością własną?

ODPOWIEDŹ: Wielokrotność $\lambda = 0$ wynosi 0.

Zadanie 4. Macierz przejścia A ma nieparzysty okres. Czy -1 jest wartością własną?

ODPOWIEDŹ: Nie. Okres musiałby być parzysty.

Zadanie 5. Przyjmijmy, że -1 jest wartością własną A . Co możemy powiedzieć o jej podległym grafie?

ODPOWIEDŹ: Jest dwudzielny.

Zadanie 6. Macierz przejścia A ma okres 6. Gdzie leżą wartości własne o module równym 1? Jakie równanie spełniają?

ODPOWIEDŹ: Na kole jednostkowym w $\mathbb{C}$. Spełniają równanie $\lambda^6 = 1$.

Twierdzenie Perrona-Frobeniusa, cz. 2

► Wartości własnej λ_1 możemy równocześnie przyporządkować ściśle dodatni lewy i prawy wektor własny.

Wniosek. Rozkład stacjonarny π jest ściśle dodatni.

Obserwacja. Prawym wektorem własnym dla λ_1 jest $\mathbf{1}$, co trywialnie wynika ze stochastyczności P .

► Jeżeli dodatkowo P jest nieokresowa (czyli $|\lambda_2| < 1$), to asymptotycznie (dla wystarczająco dużych t) mamy

$$P^t = \mathbf{1}\pi + O(t^{m_2-1}|\lambda_2|^t)$$

gdzie m_2 oznacza wielokrotność algebraiczną λ_2 .

Obserwacja. Czyli $\lim_{t \rightarrow \infty} P^t(x, y) = \pi(y)$ i szybkość „zbiegania” zależy od $|\lambda_2|$.

(np. szybko dla $|\lambda_2| = 0.4$, wolno dla $|\lambda_2| = 0.999999999999$)

► Jeśli P NIE JEST nierozkładalna i jej podległy graf składa się z $k \in \mathbb{N}$ silnie spójnych składowych (czyli każda spójna składowa jest nierozkładalna), to wielokrotność $\lambda_1 = 1$ wynosi k .

Wniosek. Unikalność rozkładu stacjonarnego π wymaga nierozkładalności P .

Obserwacja. Przypuśćmy, że $k = 3$. Łatwo znaleźć trzy rozkłady stacjonarne π_1, π_2, π_3 o rozłącznych nośnikach (każdy wsparty na jednej z trzech silnie spójnych składowych). Niech $c_1, c_2, c_3 \geq 0$ będą stałymi spełniającymi równanie $c_1 + c_2 + c_3 = 1$. Wtedy możemy stworzyć wektor $\pi_c = c_1\pi_1 + c_2\pi_2 + c_3\pi_3$ (jest to tzw. *kombinacja wypukła*), który również jest rozkładem stacjonarnym.

Część 1: Łańcuchy Markowa

Część 2: Powtórka z algebry liniowej

Część 3: Twierdzenie Perrona-Frobeniusa

Część 4: Diagonalizacja odwracalnych łańcuchów Markowa

Jeżeli dla każdej pary $x, y \in \mathcal{X}$ rozkład prawdopodobieństwa $\nu \in [0, 1]^{|\mathcal{X}|}$ spełnia

$$\nu(x)P(x, y) = \nu(y)P(y, x), \quad (*)$$

to macierz przejścia P jest **odwracalna** względem ν .

Fakt. Jeśli P jest odwracalna względem ν , to ν jest rozkładem stacjonarnym P (czyli $\nu = \pi$).

Dowód. Sumujemy obustronnie (*):

$$\sum_{x \in \mathcal{X}} \nu(x)P(x, y) = \sum_{x \in \mathcal{X}} \nu(y)P(y, x) = \nu(y) \left[\sum_{x \in \mathcal{X}} P(y, x) \right] = \nu(y). \text{ Czyli } \nu P = \nu. \quad \square$$

Fakt. Jeśli P jest odwracalna, to jej okres jest ≤ 2 .

Dowód. Warunek (*) wymusza istnienie cyklu długości 2. □

Przestrzenie $\ell^2(\pi)$ i $\ell^2(\frac{1}{\pi})$

Niech P będzie łańcuchem odwracalnym względem π . Wtedy przez $\ell^2(\pi)$ oznaczamy przestrzeń Hilberta na $\mathbb{R}^{|\mathcal{X}|}$ z iloczynem skalarnym

$$\langle f, g \rangle_\pi := \sum_{x \in \mathcal{X}} f(x)g(x)\pi(x)$$

i odpowiadającą mu normą

$$\|f\|_\pi := \sqrt{\sum_{x \in \mathcal{X}} f(x)^2 \pi(x)} = \sqrt{\langle f, f \rangle}.$$

Analogicznie definiujemy przestrzeń $\ell^2(\frac{1}{\pi})$ z $\langle f, g \rangle_{\frac{1}{\pi}}$ oraz $\|f\|_{\frac{1}{\pi}}$.

Fakt. Macierz P jest odwracalna względem π wtw macierz P jest samosprężona w $\ell^2(\pi)$, czyli dla dowolnych funkcji $f, g \in \ell^2(\pi)$ zachodzi $\langle f, Pg \rangle_\pi = \langle Pf, g \rangle_\pi$.

Diagonalizacja odwracalnych łańcuchów, cz. 1

Niech P będzie nierozkładalna i odwracalna względem π . Jej wektory i wartości własne oznaczmy W wtedy:

► Wartości własne P są rzeczywiste i mają postać:

$$1 = \lambda_1 > \lambda_2 \geq \dots \lambda_{|\mathcal{X}|-1} \geq \lambda_{|\mathcal{X}|} \geq -1.$$

► $\ell^2(\pi)$ ma bazę ortonormalną (względem $\langle \cdot, \cdot \rangle_\pi$) złożoną z prawych wektorów własnych macierzy P , czyli $f_1, f_2, \dots, f_{|\mathcal{X}|}$.

► Macierz P może być przedstawiona jako:

$$P^t = \sum_{j=1}^{|\mathcal{X}|} f_j f_j^T D_\pi \lambda_j^t = F D_\lambda^t F^T D_\pi$$

$$P^t(x, y) = \sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \pi(y) \lambda_j^t,$$

gdzie $D_\pi = \text{diag}(\pi(1), \pi(2), \dots, \pi(|\mathcal{X}|))$, $D_\lambda = \text{diag}(\lambda_1, \dots, \lambda_{|\mathcal{X}|})$

oraz $F = \begin{bmatrix} f_1 & f_2 & \dots & f_{|\mathcal{X}|} \end{bmatrix}$.

Diagonalizacja odwracalnych łańcuchów, cz. 2

► $\ell^2(\frac{1}{\pi})$ ma bazę ortonormalną (względem $\langle \cdot, \cdot \rangle_{\frac{1}{\pi}}$) złożoną z lewych wektorów własnych macierzy P , czyli $u_1, u_2, \dots, u_{|\mathcal{X}|}$.

► Macierz P może być przedstawiona jako:

$$P^t = \sum_{j=1}^{|\mathcal{X}|} D_{1/\pi} u_j u_j^T \lambda_j^t = D_{1/\pi} U D_{\lambda}^t U^T$$

$$P^t(x, y) = \sum_{j=1}^{|\mathcal{X}|} \frac{1}{\pi(x)} u_j(x) u_j(y) \lambda_j^t,$$

gdzie $D_{1/\pi} = \text{diag}(\frac{1}{\pi(1)}, \frac{1}{\pi(2)}, \dots, \frac{1}{\pi(|\mathcal{X}|)})$ oraz

$$U = \begin{bmatrix} u_1 & u_2 & \dots & u_{|\mathcal{X}|} \end{bmatrix}.$$

Obserwacja. Ponieważ

$$P^t(x, y) = \sum_{j=1}^{|\mathcal{X}|} \frac{1}{\pi(x)} u_j(x) u_j(y) \lambda_j^t = \sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \pi(y) \lambda_j^t,$$

więc zachodzi równość $u_j(x) = \pi(x) f_j(x)$.

Wniosek. Macierz P jest diagonalizowalna:

$$P^t(x, y) = \sum_{j=1}^{|\mathcal{X}|} f_j(x) u_j(y) \lambda_j^t$$

$$P^t = \sum_{j=1}^{|\mathcal{X}|} F D_{\lambda}^t U^T$$

Wniosek. Zachodzą następujące równości:

$$\langle u_i, f_j \rangle = \delta_{ij} \quad \langle f_i, f_j \rangle_\pi = \delta_{ij} \quad \langle u_i, u_j \rangle_{\frac{1}{\pi}} = \delta_{ij}.$$

Wniosek wniosku. Ponieważ $\mathbf{1} = f_1$ jest prawym wektorem własnym, więc wszystkie lewe wektory własne oprócz $u_1 = \pi^T$ sumują się do 0.

Wniosek wniosku. Ponieważ $\pi = u_1^T$ jest lewym wektorem własnym o wartości własnej 1, więc wszystkie prawe wektory własne $f_j \neq f_1$ spełniają $\langle f_j, \pi \rangle = \langle f_j, \mathbf{1} \rangle_\pi = \sum_{x \in \mathcal{X}} f_j(x) \pi(x) = 0$ (mają „wartość oczekiwaną” równą 0).

Rzeczy związane z łańcuchami Markowa:

- ▶ *Markov Chains and Mixing Times*, David A. Levin, Yuval Peres (2017)
- ▶ *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*, Pierre Brémaud (2020)

Powtórka z algebry liniowej:

- ▶ *Introduction to Stochastic Processes*, Erhan Cinlar (2013)

Dziękuję za uwagę!
wojpac1@st.amu.edu.pl