

Stochastyczne modele układów oddziałujących 2024

Notatni do wykładu

Piotr Dyszewski

2024-11-05

Contents

Sylabus	5
Skrócony plan wykładu	5
Szczegółowy plan wykładu	5
Wykład 1: Łańcuchy Markowa w czasie ciągłym	9
Podstawowe definicje	10
Wykład 2: Mocna własność Markowa	19
Czasy zatrzymania	19
Charakteryzacja	22
Wykład 3: procesy i półgrupy Feller'a	25
Proces	25
Półgrupa	27
Wykład 4: Generatory	31
Od procesu do półgrupy i generatora	32
Wykład 5: od generatora do półgrupy	35
O notacji słów kilka	35
Od generatora do procesu	37

Sylabus

Dane dotyczące przedmiotu

Nazwa przedmiotu: Stochastyczne modele systemów oddziałujących

Jednostka oferująca przedmiot: Instytut Matematyczny

Założenia: Teoria prawdopodobieństwa 2 (28-MT-S-tTPrawd2)

Strona www: <https://sites.google.com/site/piotrdyszewski/teaching/SMUO>

Forma zajęć: wykład + ćwiczenia

Punkty ECTS: 6

Skrócony plan wykładu

W trakcie wykładu poruszymy następujące zagadnienia:

- **A) Procesy Markowa:** łańcuchy Markowa w czasie ciągłym, procesy Fellera, półgrupy operatorów, generatory infinitezymalne, martyngały, rozkłady stacjonarne, ruch Browna i procesy pokrewne
- **B) Układy cząstek:** konstrukcja, ergodyczność, model głosowania, model epidemii, exclusion process

Podstawowa literatura do wykładu:

- Liggett, Thomas Milton. *Continuous time Markov processes: an introduction*. Vol. 113. American Mathematical Soc., 2010.
- Liggett, Thomas Milton, and Thomas M. Liggett. *Interacting particle systems*. Vol. 2. New York: Springer, 1985.
- Liggett, Thomas M. *Stochastic interacting systems: contact, voter and exclusion processes*. Vol. 324. Springer Science & Business Media, 2013.

Szczegółowy plan wykładu

Wstępny plan tematów poruszanych na poszczególnych wykładach:

1. Łańcuchy Markowa w czasie ciągłym
2. Procesy Fellera, półgrupy i generatory
3. Od procesu do półgrupy i generatora
4. Od generatora do procesu
5. Konstrukcje generatorów, zagadnienie martyngałowe, rozkłady stacjonarne

6. Procesy dualne
7. Zaburzenia ruchu Browna
8. Konstrukcja systemów oddziałujących
9. Ergodyczność systemów oddziałujących
10. Kilka narzędzi
11. Model głosowania: przypadek rekurencyjny
12. Model głosowania: przypadek tranzytywny
13. Model epidemii: reprezentacja graficzna i addytywność
14. Model epidemii na drzewie jednorodnym
15. Exclusion process

Efekty kształcenia

Po wykładzie student:

1. Formułuje podstawowe obiekty z zakresu teorii procesów Markowa (**A**);
2. Podaje przykłady stochastycznych systemów oddziałujących wraz z ich podstawowymi własnościami (**B**);
3. Formułuje związki między generatorami, półgrupami oraz procesami Markowa (**A**);
4. Analizuje dowody prostych twierdzeń z wykładu z uzasadnieniem poszczególnych ich założeń (**A**);
5. Formułuje główne twierdzenia teorii procesów Markowa (**A**);
6. Analizuje dowody najważniejszych twierdzeń z wykładu z uzasadnieniem poszczególnych ich założeń (**A**, **B**);
7. Formułuje główne twierdzenia z zakresu teorii systemów oddziałujących (**B**);
8. Stosuje teorię procesów Markowa w przykładach (**A**).

Sposób weryfikacji efektów kształcenia

Na zaliczenie składać się będą:

- Aktywność na ćwiczeniach;
- Zadania domowe.

Metody i kryteria oceniania

Zaliczenie ćwiczeń na podstawie zadań domowych i aktywności w czasie zajęć. Ocena z egzaminu wystawiona jest na podstawie egzaminu ustnego.

Warunkiem zaliczenia przedmiotu jest:

- Uzyskanie 30% punktów za zadania stanowiące bieżącą weryfikację efektów kształcenia;
- Uzyskanie pozytywnej oceny z egzaminu stanowiącego końcową weryfikację efektów kształcenia.

Kryteria ocen:

- (**dst**) student realizuje punkty 1-2 efektów kształcenia

- **(db)** student realizuje punkty 1-5 efektów kształcenia
 - **(bdb)** student realizuje punkty 1-8 efektów kształcenia
-

Wrocław, wrzesień 2024

Piotr Dyszewski

Wykład 1: Łańcuchy Markowa w czasie ciągłym

2024-10-03

Piotr Dyszewski

Celem tego rozdziału jest skonstruowanie procesów Markowa w czasie ciągłym na przeliczalnym (lub skończonym) zbiorze S w oparciu o jego opis infinitesimalny. W następnym rozdziale zbadamy problem konstrukcji dla procesów na bardziej ogólnej przestrzeni stanów. Na razie ograniczamy naszą uwagę do bardziej konkretnej sytuacji przeliczalnego S . W tym przypadku często używa się słowa “łańcuch” zamiast “proces”.

Przypomnijmy, że łańcuchem Markowa w czasie dyskretnym nazywamy proces stochastyczny $\{X_n\}_{n \in \mathbb{N}}$ takie, że dla dowolnego $n \in \mathbb{N}$ i dowolnych $s_0, s_1, \dots, s_n \in S$ takich, że

$$\mathbb{P}[X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2}, \dots, X_0 = s_0] > 0$$

zachodzi

$$\mathbb{P}[X_n = s_n \mid X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2}, \dots, X_0 = s_0] = \mathbb{P}[X_n = s_n \mid X_{n-1} = s_{n-1}].$$

Powyższa własność jest bardzo często przytaczana jako wyjściowa definicja łańcucha Markowa. Mimo swojej prostoty, która ułatwia czytelnikom pierwsze zetknięcie z własnością Markowa, własność ta ma jedną wadę, która ujawnia się przy bardziej zaawansowanych rozważaniach teoretycznych.

Jeżeli chcemy badać tylko procesy na **przeliczalnej** przestrzeni stanów, to jedno naturalne uogólnienie ma następującą formę. Proces stochastyczny w czasie ciągłym $\{X(t)\}_{t \in \mathbb{R}_+}$ nazwiemy łańcuchem Markowa w czasie ciągłym na przeliczalnej przestrzeni stanów S , jeżeli dla dowolnego n i dowolnych $0 \leq t_0 < t_1 < \dots < t_n$ i dowolnych $s_0, s_1, \dots, s_n \in S$ takich, że

$$\mathbb{P}[X(t_{n-1}) = s_{n-1}, X(t_{n-2}) = s_{n-2}, \dots, X(t_0) = s_0] > 0$$

zachodzi

$$\mathbb{P}[X(t_n) = s_n \mid X(t_{n-1}) = s_{n-1}, X(t_{n-2}) = s_{n-2}, \dots, X(t_0) = s_0] = \mathbb{P}[X(t_n) = s_n \mid X(t_{n-1}) = s_{n-1}].$$

Powyższa własność nie jest zbyt przydatna, jeżeli chcemy badać procesy na nieprzeliczalnej przestrzeni stanów. Dla bardzo wielu naturalnych obiektów zmienne losowe $X(t)$ w badanym przez nas procesie mogą mieć rozkład ciągły. Oznacza to, że warunek powyższy nie jest spełniony dla dowolnego wyboru parametrów.

W celu znalezienia bardziej elastycznego warunku zauważmy, że własność Markowa dla jednorodnego łańcucha Markowa $\{X_n\}_{n \in \mathbb{N}}$ w czasie dyskretnym z macierzą przejścia

$$p(i, j) = \mathbb{P}[X_1 = j \mid X_0 = i]$$

zapisuje się jako

$$\mathbb{P}[X_n = j \mid \mathcal{F}_n] = p(X_n, j),$$

gdzie $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$. Dokładne uzasadnienie powyższej własności pozostawiamy jako zadanie. Podobnie, dla dowolnego $m \in \mathbb{N}$,

$$\mathbb{P}[X_{n+m} = j \mid \mathcal{F}_n] = p^{(m)}(X_n, j),$$

gdzie $(p^{(m)(i,j)})_{i,j \in S}$ jest m -tą potęgą macierzy przejścia

$$p^{(m)}(i, j) = \mathbb{P}[X_m = j \mid X_0 = i].$$

Oznacza to, że dla dowolnej funkcji mierzalnej $f: S \rightarrow \mathbb{R}$,

$$\mathbb{E}[f(X_{n+m}) \mid \mathcal{F}_n] = \sum_{s \in S} f(s) p^{(m)}(X_n, s).$$

Powyższa definicja względnie łatwo zapisuje się w czasie ciągłym

$$\mathbb{E}[f(X_{s+t}) \mid \mathcal{F}_s] = \sum_{x \in S} f(x) p^{(t)}(X_s, x),$$

gdzie

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

Relacja powyższa daje się zapisać w przypadku nieprzeliczalnej przestrzeni stanów jako

$$\mathbb{E}[f(X_{s+t}) \mid \mathcal{F}_s] = \int_S f(x) p^{(t)}(X_s, dx),$$

gdzie

$$p^{(t)}(y, dx) = \mathbb{P}[X_t \in dx \mid X_0 = y]$$

jest rozkładem X_t pod warunkiem $\{X_0 = y\}$. To niesie ze sobą kolejne problemy, ponieważ jak wcześniej zauważyliśmy $\{X_0 = y\}$ może być zdarzeniem o prawdopodobieństwie zero. Przedstawione podejście jest do uratowania pod kątem formalnym przez odniesienie się do regularnych rozkładów warunkowych.

Zamiast tego podejźmy do problemu od innej strony. Naszym punktem wyjścia będzie odpowiednia rodzina miar. Jak zobaczymy wkrótce, to podejście będzie również opierało się o odpowiednik powyższej relacji. Oznacza to, że w rezultacie będziemy opisywali tę samą klasę procesów stochastycznych bez konieczności obchodzenia się z warunkowaniem po zdarzeniach niemożliwych.

Podstawowe definicje

Zaczynamy od prezentacji definicji trzech obiektów, na których się skupimy w tym rozdziale. Przypomnijmy, że będziemy definiować łańcuchy Markowa w czasie ciągłym na dyskretnej przestrzeni

stanów S . Topologia na S to oczywiście topologia dyskretna, względem której wszystkie funkcje są ciągłe.

Głównym obiektem naszych badań będą procesy stochastyczne. Ze względów technicznych pracować będziemy na bardzo konkretnej przestrzeni zdarzeń elementarnych. Niech Ω będzie zbiorem prawostronnie ciągłych funkcji $\omega: [0, \infty) \rightarrow S$ ze skończoną liczbą skoków w dowolnym skończonym przedziale czasowym. Dla każdego $t \in \mathbb{R}_+ = [0, +\infty)$ rozważmy funkcję $X_t: \Omega \rightarrow S$ zadaną przez

$$X_t(\omega) = \omega(t).$$

Niech σ -ciało \mathcal{F} na Ω będzie najmniejszym takim, że odwzorowanie $\omega \mapsto \omega(t)$ jest mieralne dla każdego $t \in \mathbb{R}_+ = [0, +\infty)$. Niech wreszcie dla $s \in \mathbb{R}_+$ oznaczmy przez θ_s odwzorowanie $\Omega \rightarrow \Omega$ zadane przez

$$\theta_s(\omega)(t) = \omega(t + s).$$

W szczególności $X_t \circ \theta_s = X_{t+s}$. O odwzorowaniu θ_s można myśleć jak o przesunięciu czasu o s . Zamiast utożsamiać własność Markowa z procesem stochastycznym $X = \{X(t)\}_{t \in \mathbb{R}_+}$, utożsamimy ją z rodziną miar probabilistycznych na Ω , względem której X będzie procesem Markowa.

Definicja 0.1. Łańcuchem Markowa w czasie ciągłym na przestrzeni stanów S nazywamy parę uporządkowaną (\mathbf{P}, \mathbb{F}) taką, że

- **(ŁM1)** $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$ jest filtracją względem której $X = (X(t))_{t \in \mathbb{R}_+}$ jest adaptowalny i $\mathcal{F}_t \subseteq \mathcal{F}$ dla każdego $t \in \mathbb{R}_+$.
- **(ŁM2)** $\mathbf{P} = \{\mathbf{P}_x\}_{x \in S}$. Dla każdego $x \in S$, \mathbf{P}_x jest miarą probabilistyczną na Ω taką, że

$$\mathbf{P}_x[X_0 = x] = \mathbf{P}_x[\omega \in \Omega : \omega(0) = x] = 1.$$

- **(ŁM3)** Spełniona jest własność Markowa

$$\mathbf{E}_x[Y \circ \theta_s | \mathcal{F}_s] = \mathbf{E}_{X(s)}[Y] \text{ p.n. } \mathbf{P}_x$$

dla wszystkich $x \in S$ i wszystkich ograniczonych mierzalnych Y na Ω .

W powyższej definicji \mathbf{E}_x jest wartością oczekiwaną odpowiadającą mierze probabilistycznej \mathbf{P}_x , czyli

$$\mathbf{E}_x[Y] = \int_{\Omega} Y(\omega) \mathbf{P}_x(d\omega).$$

O zmiennej losowej $Y: \Omega \rightarrow \mathbb{R}$ można myśleć jak o statystyce całej trajektorii procesu $\{X(t)\}_{t \in \mathbb{R}_+}$. Załóżmy, że Y jest postaci

$$Y(\omega) = f(\omega(t_1), \omega(t_2), \dots, \omega(t_n)) = f(X(t_1), X(t_2), \dots, X(t_n))$$

dla pewnej mierzalnej i ograniczonej funkcji $f: \mathbb{R}^n \rightarrow \mathbb{R}$ (dla ćwiczenia warto sprawdzić, że Y powyższej postaci jest istotnie zmienną losową, tj. jest mierzalna względem \mathcal{F}). Wówczas

$$Y \circ \theta_s(\omega) = f(\omega(t_1 + s), \omega(t_2 + s), \dots, \omega(t_n + s)) = f(X(t_1 + s), X(t_2 + s), \dots, X(t_n + s)).$$

Własność Markowa w Definicji mówi zatem to, co powinna. Rozkład wektora losowego $(X(t_1 + s), X(t_2 + s), \dots, X(t_n + s))$ pod warunkiem \mathcal{F}_s jest taki sam, jak rozkład wektora $(X'(t_1), X'(t_2), \dots, X'(t_n))$ dla $X' = \{X'(t)\}_{t \in \mathbb{R}_+}$ będącym niezależną kopią X , zapoczątkowaną w $X'_0 = X(s)$.

Naszym nadrzędnym celem będzie sprowadzenie powyższej definicji do bardziej przystępnych terminów. Zanim jednak do tego przejdziemy, rozważmy następujący przykład.

Przykład 0.1. Niech $N = (N_t)_{t \in \mathbb{R}_+}$ będzie jednorodnym procesem Poissona z intensywnością $\lambda > 0$ określonym na przestrzeni probabilistycznej $(\Sigma, \mathcal{G}, \mathbb{P})$. Przypomnijmy, że oznacza to, że $t \mapsto N_t$ jest prawostronnie ciągła oraz dla każdego $t > 0$ zmienna losowa N_t ma rozkład Poissona z parametrem $\lambda > 0$,

$$\mathbb{P}[N_t = k] = e^{-\lambda} \frac{\lambda^k}{k!},$$

i wreszcie dla $t > s \geq 0$, zmienna $N_t - N_s$ jest niezależna od σ -ciała $\mathcal{F}_s^N = \sigma(N_r : r \leq s)$.

Uzasadnimy, że N jest procesem Markowa w sensie przyjętej przez nas definicji. Przestrzenią stanów jest $S = \mathbb{N}$. Niech $\mathbf{P}_x[A] = \mathbb{P}[N + x \in A]$ dla $A \in \mathcal{F}$. Tutaj $x + N$ oznacza funkcję $t \mapsto N_t + x$. Pokażemy teraz, że spełniona jest własność Markowa.

Dla $A \in \mathcal{F}_s$ mamy

$$\mathbf{E}_x[Y \circ \theta_s \cdot \mathbf{1}_A] = \mathbb{E}[Y(N \circ \theta_s + x) \mathbf{1}_{\{N \in A\}}] = \mathbb{E}[Y(N \circ \theta_s - N_s + N_s + x) \mathbf{1}_{\{N \in A\}}].$$

Skoro $A \in \mathcal{F}_s$, to $\{N \in A\} \in \mathcal{F}_s^N$ (zadanie). Skoro $N \circ \theta_s - N_s = (N_{t+s} - N_s)_{t \in \mathbb{R}_+}$ jest niezależny od \mathcal{F}_s^N , to

$$\mathbb{E}[Y(N \circ \theta_s - N_s + N_s + x) \mathbf{1}_{\{N \in A\}} | \mathcal{F}_s] = \mathbf{1}_{\{N \in A\}} \cdot y(N_s + x),$$

gdzie

$$y(k) = \mathbb{E}[Y(N \circ \theta_s - N_s + k)] = \mathbf{E}_k[Y].$$

Podsumowując,

$$\mathbf{E}_x[Y \circ \theta_s \cdot \mathbf{1}_A] = \mathbb{E}[y(N_s + x) \mathbf{1}_{\{N \in A\}}] = \mathbf{E}_x[y(X(s)) \mathbf{1}_A]$$

odwołując się teraz do definicji warunkowej wartości oczekiwanej

$$\mathbf{E}_x[Y \circ \theta_s | \mathcal{F}_s] = y(X(s)) = \mathbf{E}_{X(s)}[Y].$$

Powyższy przykład pokazuje, że uzasadnienie własności Markowa wprost z definicji nie jest najprostszym zadaniem. Przekonamy się w przyszłości, że taka forma jest przydatna do teoretycznych rozważań. Mimo to przyda się nam bardziej przystępny sposób mówienia o łańcuchach Markowa w czasie ciągłym. Odnosząc się do czasu dyskretnego, korzystać będziemy z funkcji przejścia.

Definicja 0.2. Funkcją przejścia nazywamy rodzinę odwzorowań $p = (p_t)_{t \in \mathbb{R}_+}$, gdzie $p_t : S \times S \rightarrow [0, 1]$ zdefiniowanych dla $t \geq 0$ takich, że

$$p_t(x, y) \geq 0, \quad \sum_{y \in S} p_t(x, y) = 1, \quad \lim_{t \rightarrow 0} p_t(x, x) = p_0(x, x) = 1,$$

spełniających równania Chapmana-Kolmogorowa

$$p_{s+t}(x, y) = \sum_{z \in S} p_s(x, z) p_t(z, y).$$

Interpretacją wartości $p_t(x, y)$ jest prawdopodobieństwo, że w czasie t proces przejdzie ze stanu x do stanu y . Innymi słowy,

$$p_t(x, y) = \mathbf{P}_x[X_t = y].$$

Jak się niebawem przekonamy, dzięki własności Markowa pozwala ona jednoznacznie wyznaczyć rozkład procesu, tj. jednoznacznie wyznaczyć miarę \mathbf{P}_x .

Przykład 0.2. Rozważmy $S = \mathbb{N}$, $\lambda > 0$ oraz $p_t: S \times S \rightarrow [0, 1]$ zadane przez

$$p_t(x, y) = e^{-\lambda t} \frac{(\lambda t)^{y-x}}{(y-x)!} \mathbf{1}_{\{y \geq x\}}.$$

Wówczas p jest funkcją przejścia. Wystarczy zauważyć, że $p_t(x)$ to prawdopodobieństwo, że zmienna losowa o rozkładzie Poissona z parametrem λt jest równa $y - x$. Równania Chapmana-Kołmogorowa wynikają z następującej własności rozkładu Poissona: jeżeli niezależne zmienne losowe X i Y mają rozkłady Poissona odpowiednio z parametrami λt i λs , to $X + Y$ ma rozkład Poissona z parametrem $\lambda(t + s)$.

Twierdzenie 0.1. Dla łańcucha Markowa (\mathbf{P}, \mathbb{F}) połóżmy

$$p_t(x, y) = \mathbf{P}_x[X_t = y]$$

dla $t \geq 0$ oraz $x, y \in S$. Wówczas:

- a. $p_t(x, y)$ jest funkcją przejścia,
- b. $p_t(x, y)$ określa miary \mathbf{P}_x jednoznacznie.

Proof.

- a. Najpierw pokażemy, że $\lim_{t \rightarrow 0} p_t(x, x) = 1$. Przez prawostronną ciągłość ścieżek, $\tau = \inf\{t > 0 : X_t \neq X_0\} > 0$ \mathbf{P}_x -p.w. dla dowolnego x z S . Ponieważ

$$p_t(x, x) \geq \mathbf{P}_x[\tau > t] \rightarrow 1$$

przy $t \rightarrow 0$, to istotnie $\lim_{t \rightarrow 0} p_t(x, x) = 1$.

Równania Chapmana-Kołmogorowa wynikają z własności Markowa. Aby to zobaczyć, rozważmy $Y = 1_{\{X(t)=y\}}$. Własność Markowa zapisuje się jako

$$\mathbf{P}_x[X_{s+t} = y | \mathcal{F}_s] = \mathbf{P}_{X(s)}[X_t = y] = p_t(X(s), y) \text{ p.n. } \mathbf{P}_x.$$

Biorąc wartości oczekiwane względem \mathbf{P}_x w tej tożsamości, otrzymujemy równania Chapmana-Kołmogorowa:

$$\mathbf{P}_x[X_{t+s} = y] = \mathbf{E}_x[\mathbf{P}_x[X_{s+t} = y | \mathcal{F}_s]] = \mathbf{E}_x[p_t(X(s), y)] = \sum_{z \in S} p_t(z, y) p_s(x, z).$$

To dowodzi a. b. Użyjemy własności Markowa wielokrotnie, aby otrzymać

$$\mathbf{P}_x[X(t_1) = x_1, \dots, X(t_n) = x_n] = p_{t_1}(x, x_1) p_{t_2-t_1}(x_1, x_2) \cdots p_{t_n-t_{n-1}}(x_{n-1}, x_n)$$

dla $0 < t_1 < \dots < t_n$ oraz $x_1, \dots, x_n \in S$. Aby to zobaczyć oznaczmy

$$\mathcal{H}_{n-1} = \{X(t_1) = x_1, \dots, X(t_{n-1}) = x_{n-1}\} \in \mathcal{F}_{t_{n-1}}.$$

Z własności Markowa

$$\begin{aligned} \mathbf{P}_x[X(t_1) = x_1, \dots, X(t_n) = x_n | \mathcal{F}_{t_{n-1}}] &= \mathbf{P}_x[\mathcal{H}_{n-1}, X(t_n) = x_n | \mathcal{F}_{t_{n-1}}] \\ &= \mathbf{1}_{\mathcal{H}_{n-1}} \mathbf{P}_x[X(t_n) = x_n | \mathcal{F}_{t_{n-1}}] = \mathbf{1}_{\mathcal{H}_{n-1}} \mathbf{P}_x[X \circ \theta_{t_{n-1}}(t_n - t_{n-1}) = x_n | \mathcal{F}_{t_{n-1}}] \end{aligned}$$

$$= \mathbf{1}_{\mathcal{H}_{n-1}} \mathbf{P}_{X(t_{n-1})}[X(t_n - t_{n-1}) = x_n] = \mathbf{1}_{\mathcal{H}_{n-1}} p_{t_n - t_{n-1}}(x_{n-1}, x_n).$$

Biorąc wartości oczekiwane,

$$\mathbf{P}_x[X(t_1) = x_1, \dots, X(t_n) = x_n] = \mathbf{P}_x[\mathcal{H}_{n-1}] p_{t_n - t_{n-1}}(x_{n-1}, x_n).$$

Postulowaną równość otrzymujemy przez iterację powyższej procedury. Udowodniona właśnie równość uzasadnia, że funkcja przejścia określa rozkłady skończenie wymiarowe \mathbf{P}_x . Określa również pełną miarę \mathbf{P}_x , ponieważ miary prawdopodobieństwa na (Ω, \mathcal{F}) są określane przez ich skończenie wymiarowe rozkłady w świetle twierdzenia $\pi - \lambda$. Aby to zobaczyć, załóżmy, że μ oraz ν to dwie takie miary, które mają te same skończenie wymiarowe rozkłady, i niech \mathcal{P} będą skończenie wymiarowymi zbiorami w (Ω, \mathcal{F}) oraz

$$\mathcal{L} = \{A \in \mathcal{F} : \mu(A) = \nu(A)\}.$$

Wówczas \mathcal{L} jest λ -układem zawierającym π -układ \mathcal{P} . Przez twierdzenie $\pi - \lambda$, $\sigma(\mathcal{P}) \subseteq \mathcal{L}$. Skoro $\sigma(\mathcal{P}) = \mathcal{F}$, to $\mu(A) = \nu(A)$ dla wszystkich $A \in \mathcal{F}$.

□

Skoro X jest elementem Ω , zbioru funkcji prawostronnie ciągłych o wartościach w przeliczalnym S , to X jest funkcją kawałkami stałą. W rezultacie do opisu X wystarczy sprecyzować, w jaki sposób X zmienia wartość. Opis ten jest dokonywany w kategoriach Q -macierzy.

Definicja 0.3. Q -macierzą nazywamy macierz $(q(x, y))_{x, y \in S}$ liczb rzeczywistych indeksowanych przez $x, y \in S$, które spełniają

$$q(x, y) \geq 0 \text{ dla } x \neq y \quad \text{oraz} \quad \sum_y q(x, y) = 0.$$

Ponieważ wyrazy diagonalne są niedodatnie i odgrywają specjalną rolę, naturalne jest użycie specjalnego oznaczenia dla nich:

$$c(x) = -q(x, x).$$

Przejście od funkcji przejścia do Q -macierzy jest trudniejsze niż przejście od procesu do funkcji przejścia i wymaga dokonania dodatkowych założeń. Zaczynamy od kilku własności, które obowiązują dla wszystkich funkcji przejścia.

Twierdzenie 0.2. *Założmy, że p jest funkcją przejścia.*

- a. Wówczas $p_t(x, x) > 0$ dla wszystkich $t \geq 0$ oraz $x \in S$.
- b. Jeśli $p_t(x, x) = 1$ dla pewnego $t > 0$ oraz $x \in S$, wtedy $p_t(x, x) = 1$ dla wszystkich $t > 0$ oraz tego x .
- c. Dla każdego $x, y \in S$, $p_t(x, y)$ jest jednostajnie ciągła w t . Dokładniej,

$$|p_t(x, y) - p_s(x, y)| \leq 1 - p_{|t-s|}(x, x).$$

Proof.

- a. Najpierw zauważmy, że $p_t(x, x) > 0$ dla małych t zgodnie z przyjętą Definicją 0.2 Z równań Chapmana-Kołmogorowa,

$$p_{s+t}(x, x) \geq p_s(x, x)p_t(x, x),$$

więc ścisła dodatniość rozciąga się na wszystkie t .

b. Użyjmy równania Chapmana-Kołmogorowa raz jeszcze, aby napisać

$$p_{s+t}(x, x) \leq p_s(x, x)p_t(x, x) + [1 - p_s(x, x)] = 1 - p_s(x, x)[1 - p_t(x, x)].$$

Zatem, jeśli $p_{s+t}(x, x) = 1$, to $p_t(x, x) = 1$, ponieważ $p_s(x, x) > 0$ z części a. Stąd $\{t \geq 0 : p_t(x, x) = 1\}$ jest przedziałem zaczynającym się od 0. Z dowodu części a. wynika, że musi to być cała dodatnia oś.

c. Ponownie użyjmy równania Chapmana-Kołmogorowa, aby napisać

$$p_{t+s}(x, y) - p_t(x, y) = p_t(x, y)[p_s(x, x) - 1] + \sum_{z \neq x} p_s(x, z)p_t(z, y).$$

Pierwszy składnik po prawej stronie jest niedodatni, a drugi jest nieujemny. Wartość bezwzględna każdego z nich nie jest większa niż $1 - p_s(x, x)$. To pociąga postulowaną nierówność, która z kolei pociąga jednostajną ciągłość.

□

Twierdzenie 0.3. *Załóżmy, że p jest funkcją przejścia.*

a. *Dla każdego x , prawostronna pochodna*

$$c(x) = -q(x, x) = - \left. \frac{d}{dt} p_t(x, x) \right|_{t=0} \in [0, \infty]$$

istnieje i spełnia

$$p_t(x, x) \geq e^{-c(x)t}.$$

b. *Jeśli $c(x) < \infty$, wtedy dla tego x i dla wszystkich $y \neq x$, prawostronna pochodna*

$$q(x, y) = \left. \frac{d}{dt} p_t(x, y) \right|_{t=0} \in [0, \infty)$$

istnieje oraz

$$\sum_{y \in S} q(x, y) \leq 0.$$

c. *Jeśli dla pewnego $x \in S$, $c(x) < \infty$ i $\sum_y q(x, y) = 0$, to $p_t(x, y)$ jest różniczkowalna w sposób ciągły względem t dla tego x i każdego y , oraz spełnia równania retrospektywne Kołmogorowa:*

$$\frac{d}{dt} p_t(x, y) = \sum_z q(x, z) p_t(z, y).$$

Proof.

a. Niech $f(t) = -\log p_t(x, x)$. Wówczas f jest dobrze określona i jednostajnie ciągła. Jest ona też subaddytywna. Zatem, zgodnie z Lematem Fekete,

$$c(x) = \lim_{t \rightarrow 0} \frac{f(t)}{t} \in [0, \infty]$$

i spełnia $f(t) \leq c(x)t$.

b. Przypuśćmy, że $c(x) < \infty$. Z dowiedzionej właśnie nierówności,

$$1 - p_t(x, x) \leq 1 - e^{-c(x)t} \leq c(x)t,$$

i stąd

$$\sum_{y: y \neq x} \frac{p_t(x, y)}{t} \leq c(x).$$

Zatem,

$$\limsup_{t \rightarrow 0} \frac{p_t(x, y)}{t} < \infty$$

dla $y \neq x$. Niech $q(x, y)$ będzie wartością powyższej granicy górnej. Aby pokazać, że granica rzeczywiście istnieje, weźmy $\delta > 0$ i dodatnią liczbę całkowitą n . Macierz $p_\delta(x, y)$ można traktować jako prawdopodobieństwa przejścia dla dyskretnego łańcucha Markowa na S , a wtedy zgodnie z równaniami Chapmana-Kołmogorowa, odpowiadające n -krokowe prawdopodobieństwa przejścia są dane przez $p_{n\delta}(x, y)$. Rozkładając zdarzenie, że ten łańcuch znajduje się w y w czasie n zgodnie z czasem pierwszej wizyty w y , mamy dla $y \neq x$,

$$p_{n\delta}(x, y) \geq \sum_{k=0}^{n-1} p_{k\delta}^k(x, x) p_\delta(x, y) p_{(n-k-1)\delta}(y, y).$$

Stąd

$$\frac{p_{n\delta}(x, y)}{n\delta} \geq \frac{p_\delta(x, y)}{\delta} e^{-c(x)n\delta} \inf_{0 \leq s \leq n\delta} p_s(y, y).$$

Teraz niech $\delta \downarrow 0$ wzdłuż ciągu realizującego granicę górną w (2), tak że

$$\frac{p_\delta(x, y)}{\delta} \rightarrow q(x, y).$$

Wyberzmy teraz $n \rightarrow \infty$ tak, że $n\delta \rightarrow t$. Wówczas

$$\frac{p_t(x, y)}{t} \geq q(x, y) e^{-c(x)t} \inf_{0 \leq s \leq t} p_s(y, y)$$

dla $t > 0$. Zatem,

$$\liminf_{t \rightarrow 0} \frac{p_t(x, y)}{t} \geq q(x, y).$$

Postulowana nierówność wynika teraz z lematu Fatou.

c. Napiszmy

$$\frac{p_{t+s}(x, y) - p_t(x, y)}{s} = \sum_z \left[\frac{p_s(x, z) - p_0(x, z)}{s} - q(x, z) \right] p_t(z, y).$$

Każdy wyraz w sumie dąży do 0, gdy $s \downarrow 0$, zgodnie z pierwszymi dwoma częściami twierdzenia. Zatem musimy kontrolować ogony sumy. Weźmy skończony zbiór $T \subset S$ zawierający x i zauważmy, że

$$\sum_{z \notin T} \left| \frac{p_s(x, z)}{s} - q(x, z) \right| p_t(z, y) \leq \sum_{z \notin T} \frac{p_s(x, z)}{s} + \sum_{z \notin T} q(x, z)$$

$$= s^{-1} \left[1 - \sum_{z \in T} p_s(x, z) \right] - \sum_{z \in T} q(x, z) \rightarrow -2 \sum_{z \in T} q(x, z)$$

gdy $s \downarrow 0$. Granicę po prawej stronie można uczynić dowolnie małą, wybierając T duże, ponieważ $\sum_z q(x, z) = 0$. Zatem prawa strona dąży do zera, gdy $s \downarrow 0$. To dowodzi, że prawostronne pochodne $p_t(x, y)$ istnieją i spełniają zadaną równość. Aby zobaczyć, że obustronne pochodne rzeczywiście istnieją, wystarczy zauważyć, że prawa strona jest ciągła w t i użyć faktu, że funkcja ciągła z ciągłą prawostronną pochodną jest różniczkowalna.

□

Wykład 2: Mocna własność Markowa

2024-10-10

Piotr Dyszewski

Pojęcie czasu zatrzymania odgrywa kluczową rolę w teorii procesów stochastycznych. Są to losowe momenty adaptowalne do z góry zadanej filtracji. Jest to kluczowa koncepcja w silnej własności Markowa. Będziemy korzystać z ciągłej filtracji $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$.

Czasy zatrzymania

Przypomnijmy, że w czasie dyskretnym definicja czasu zatrzymania τ jest taka, że $\{\tau = n\} \in \mathcal{F}_n$ dla każdego naturalnego n . Jest to równoważne z warunkiem, że $\{\tau \leq n\} \in \mathcal{F}_n$ dla każdego naturalnego n . W czasie ciągłym ta równoważność nie zachodzi, ponieważ $[0, \infty)$ nie jest przeliczalne. Warunek analogiczny do tego drugiego jest naturalny do użycia w czasie ciągłym, ponieważ zazwyczaj zdarzenie $\{\tau = t\}$ ma zerowe prawdopodobieństwo dla każdego t .

Definicja 0.4. Zmienna losowa $\tau : \Omega \rightarrow [0, \infty]$ nazywana jest \mathbb{F} -czasem zatrzymania, jeśli $\{\tau \leq t\} \in \mathcal{F}_t$ dla każdego $t \geq 0$.

W niektórych kontekstach filtracja \mathbb{F} , z którą pracujemy, jest na tyle regularna, że ułatwia to weryfikację, czy zmienna jest czasem zatrzymania.

Definicja 0.5. Powiemy, że filtracja $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{R}_+}$ jest prawostronnie ciągła, jeżeli

$$\mathcal{F}_t = \mathcal{F}_{t+}, \quad \text{gdzie} \quad \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$$

dla każdego $t \in \mathbb{R}_+$.

Zadanie 0.1. Załóżmy, że filtracja \mathbb{F} jest prawostronnie ciągła. Wówczas τ jest czasem zatrzymania wtedy i tylko wtedy, gdy $\{\tau < t\} \in \mathcal{F}_t$ dla każdego $t \in \mathbb{R}_+$.

Zadanie 0.2. Pokaż, że jeśli τ_1 i τ_2 są czasami zatrzymania, to również $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$ i $\tau_1 + \tau_2$ są czasami zatrzymania.

Zadanie 0.3. Udowodnij, że jeśli $\{\tau_n\}_{n \in \mathbb{N}}$ jest ciągiem czasów zatrzymania, które maleją do τ , to τ jest czasem zatrzymania.

Własność Markowa dotyczy warunkowej wartości oczekiwanej względem \mathcal{F}_s dla ustalonego s .

Mocna własność Markowa jest analogiczna, ale warunkowanie odbywa się względem σ -algebry \mathcal{F}_τ , gdzie τ jest czasem stopu. Składa się ona ze zdarzeń, które są określone przez przeszłość aż do czasu τ .

Definicja 0.6. Dla czasu zatrzymania τ kładziemy

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ dla każdego } t \in \mathbb{R}_+\}.$$

Zadanie 0.4. Pokaż, że:

- a. \mathcal{F}_τ jest σ -algebrą,
- b. Załóżmy, że \mathbb{F} jest prawostronnie ciągła. Pokaż, że

$$\mathcal{F}_\tau = \{A : A \cap \{\tau < t\} \in \mathcal{F}_t \text{ dla każdego } t \geq 0\}.$$

Oto niektóre podstawowe własności \mathcal{F}_τ .

Twierdzenie 0.4. Jeśli $\tau, \tau_n, n \in \mathbb{N}$ są czasami zatrzymania, to:

- a. τ jest mierzalny względem \mathcal{F}_τ .
- b. Jeśli $\tau_n \downarrow \tau$, to $\mathcal{F}_\tau = \bigcap_n \mathcal{F}_{\tau_n}$.
- c. $\tau_1 \leq \tau_2$ implikuje $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

Proof. Zadanie. □

Mocna własność Markowa

Twierdzenie 0.5. Niech (\mathbb{P}, \mathbb{F}) będzie łańcuchem Markowa na przeliczalnej przestrzeni stanów S . Załóżmy, że Y jest ograniczoną zmienną losową oraz że τ jest czasem zatrzymania. Wówczas dla każdego $x \in S$,

$$\mathbf{E}_x[Y \circ \theta_\tau | \mathcal{F}_\tau] = \mathbf{E}_{X(\tau)}[Y] \quad \mathbf{P}_x - \text{prawie na pewno na } \{\tau < \infty\}. \quad (1)$$

Silna własność Markowa jest zazwyczaj używana w następujący sposób: Przemnóż równość (1) przez $\mathbf{1}_{\{\tau < \infty\}}$, a następnie zastosuj \mathbf{E}_x . Wynik, z uwzględnieniem, że $\{\tau < \infty\} \in \mathcal{F}_\tau$, to:

$$\mathbf{E}_x[Y \circ \theta_\tau \mathbf{1}_{\{\tau < \infty\}}] = \mathbf{E}_x[\mathbf{E}_{X(\tau)}[Y] \mathbf{1}_{\{\tau < \infty\}}].$$

Twierdzenia 0.5. Najpierw załóżmy, że τ przyjmuje wartości z przeliczalnego zbioru $0 \leq t_1 \leq t_2 \leq \dots$ oraz ∞ . W tym przypadku silna własność Markowa sprowadza się do własności Markowa, jak teraz pokażemy.

Zauważmy, że prawa strona (1) jest mierzalna względem \mathcal{F}_τ . Musimy więc sprawdzić, że jeśli $A \in \mathcal{F}_\tau$ oraz $A \subseteq \{\tau < \infty\}$, to

$$\mathbf{E}_x[Y \circ \theta_\tau \mathbf{1}_A] = \mathbf{E}_x[\mathbf{E}_{X(\tau)}[Y] \mathbf{1}_A].$$

Aby to wykazać, napiszmy:

$$\mathbf{E}_x[Y \circ \theta_\tau \mathbf{1}_A] = \sum_{n \in \mathbb{N}} \mathbf{E}_x[Y \circ \theta_{t_n} \mathbf{1}_{A \cap \{\tau = t_n\}}] = \sum_{n \in \mathbb{N}} \mathbf{E}_x[\mathbf{E}_{X(t_n)}[Y] \mathbf{1}_{A \cap \{\tau = t_n\}}] = \mathbf{E}_x[\mathbf{E}_{X(\tau)}[Y] \mathbf{1}_A].$$

W drugim kroku skorzystaliśmy z własności Markowa, ponieważ

$$A \cap \{\tau = t_n\} \in \mathcal{F}_{t_n}$$

zgodnie z definicją \mathcal{F}_τ .

W drugim kroku uzasadnimy tezę dla dowolnych τ i Y postaci

$$Y(\omega) = \prod_{j=1}^m f_j(\omega(t_j)), \quad (2)$$

dla pewnego $m \in \mathbb{N}$, $t_1, \dots, t_m \in \mathbb{R}_+$ oraz ograniczonych funkcji $f_1, \dots, f_m: S \rightarrow \mathbb{R}$.

Czas τ przybliżamy go od góry czasami stopu $\{\tau_n\}_{n \in \mathbb{N}}$ zdefiniowanymi przez

$$\tau_n = \frac{k+1}{2^n} \quad \text{jeśli} \quad \frac{k}{2^n} \leq \tau < \frac{k+1}{2^n}.$$

dla dostatecznie dużych k . Weźmy teraz $A \in \mathcal{F}_\tau \subseteq \mathcal{F}_{\tau_k}$ takie, że $A \subseteq \{\tau < \infty\}$. Z pierwszej części dowodu,

$$\mathbf{E}_x[Y \circ \theta_{\tau_k} \mathbf{1}_A] = \mathbf{E}_x[\mathbf{E}_{X(\tau_k)}[Y] \mathbf{1}_A].$$

Musimy przejść do granicy, gdy $k \rightarrow \infty$. Po prawej stronie, $\tau_k \downarrow \tau$ i z prawostronnej ciągłości $X(\tau_k) \rightarrow X(\tau)$ w S , czyli $X(\tau_k) = X(\tau)$ dla dostatecznie dużych k . Po lewej stronie, napiszmy

$$(Y \circ \theta_{\tau_k})(\omega) = \prod_{m=1}^n f_m(\omega(t_m + \tau_k)) \rightarrow \prod_{m=1}^n f_m(\omega(t_m + \tau)) = (Y \circ \theta_\tau)(\omega).$$

kiedy $k \rightarrow \infty$, dzięki prawostronnej ciągłości ścieżek. To pokazuje tezę:

$$\mathbf{E}_x[Y \circ \theta_\tau \mathbf{1}_A] = \mathbf{E}_x[\mathbf{E}_{X(\tau)}[Y] \mathbf{1}_A].$$

W ostatnim kroku dowodu pokażemy tezę dla dowolnego Y . Dla dowolnego $m \in \mathbb{N}$ oraz $t_1, \dots, t_m \in \mathbb{R}_+$ rozważmy $\pi_{t_1, \dots, t_m}: \Omega \rightarrow \mathbb{R}^m$ dane wzorem

$$\pi_{t_1, \dots, t_m}(\omega) = (\omega(t_1), \omega(t_2), \dots, \omega(t_m)).$$

Wówczas dla dowolnych $A_1, A_2, \dots, A_m \subseteq S$,

$$\pi_{t_1, \dots, t_m}^{-1}[A_1 \times A_2 \times \dots \times A_m] = \{\omega \in \Omega : \omega(t_1) \in A_1, \dots, \omega(t_m) \in A_m\}.$$

Rozważmy teraz π -układ

$$\mathcal{B} = \left\{ \pi_{t_1, \dots, t_m}^{-1}[A_1 \times A_2 \times \dots \times A_m] : m \in \mathbb{N}, t_1, \dots, t_m \in \mathbb{R}_+, A_1, \dots, A_m \subseteq S \right\}.$$

oraz λ -układ

$$\mathcal{L} = \left\{ G \in \mathcal{F} : \mathbf{P}_x[\theta_\tau \in G, A] = \mathbf{E}_x[\mathbf{P}_{X(\tau)}[G] \mathbf{1}_A] \text{ dla } A \in \mathcal{F}_\tau, A \subseteq \{\tau < \infty\} \right\}.$$

Aproksymując $\mathbf{1}_{\pi_{t_1, \dots, t_m}^{-1}[A_1 \times \dots \times A_m]}$ zmiennymi Y postaci (2), dostajemy $\mathcal{B} \subseteq \mathcal{L}$. Z lematu o π - λ układach mamy $\mathcal{F} = \sigma(\mathcal{B}) \subseteq \mathcal{L}$. Czyli dla każdego $G \in \mathcal{F}$,

$$\mathbf{P}_x[\theta_\tau \in G, A] = \mathbf{E}_x[\mathbf{P}_{X(\tau)}[G] \mathbf{1}_A]$$

dla każdego $A \in \mathcal{F}_\tau$ takiego, że $A \subseteq \{\tau < \infty\}$. Jest to równoważne naszej tezie dla $Y = \mathbf{1}_G$. Z liniowości teza jest zatem prawdziwa dla każdego Y przyjmującego skończenie wiele wartości. Zastosowanie standardowego twierdzenia granicznego dowodzi tezy dla dowolnego ograniczonego Y . \square

Charakteryzacja

Zauważmy, że każda funkcja $\omega \in \Omega$ musi być następującego typu: Istnieje $t_1 \in (0, \infty]$, taki że $\omega(t) = \omega(0)$ dla każdego $t \in [0, t_1)$, następnie, jeśli $t_1 < \infty$, istnieje $t_2 \in (t_1, \infty]$ taki, że $\omega(t) = \omega(t_1) \neq \omega(0)$ dla każdego $t \in [t_1, t_2)$, i tak dalej. Powyższe czasy t_1, t_2, \dots zależą oczywiście od wyboru ω . Dla każdego $\omega \in \Omega$, istnieje zatem ciąg

$$T_0(\omega) = 0 < T_1(\omega) \leq T_2(\omega) \leq T_3(\omega) \leq \dots \leq \infty,$$

taki, że $X_t(\omega) = X_0(\omega)$ dla każdego $t \in [0, T_1(\omega))$ oraz dla każdej liczby całkowitej $i \geq 1$, warunek $T_i(\omega) < \infty$ implikuje $T_i(\omega) < T_{i+1}(\omega)$, $X_{T_i(\omega)}(\omega) \neq X_{T_{i-1}(\omega)}(\omega)$ i $X_t(\omega) = X_{T_i(\omega)}(\omega)$ dla każdego $t \in [T_i(\omega), T_{i+1}(\omega))$. Co więcej, $T_n(\omega) \uparrow \infty$, gdy $n \rightarrow \infty$. Nietrudno jest sprawdzić, że T_0, T_1, T_2, \dots są czasami stopu. Na przykład,

$$\{T_1 \leq t\} = \{X(t) \neq X(0)\} \cup \bigcup_{q \in (0,1) \cap \mathbb{Q}} \{X_q \neq X_0\} \in \mathcal{F}_t.$$

Przypomnijmy, że dla $\lambda > 0$, dodatnia zmienna losowa U ma rozkład wykładniczy z parametrem λ , jeśli $\mathbb{P}[U > r] = e^{-\lambda r}$ dla każdego $r \geq 0$. W poniższym lemacie przyjmujemy konwencję, że zmienna losowa wykładnicza o parametrze 0 jest równa ∞ prawie na pewno.

Lemma 0.1. *Niech $x \in S$. Istnieje rzeczywista liczba $c(x) \geq 0$, taka że zmienna losowa T_1 ma rozkład wykładniczy z parametrem $c(x)$ pod \mathbf{P}_x . Co więcej, jeśli $c(x) > 0$, to T_1 i X_{T_1} są niezależne pod \mathbf{P}_x .*

Proof. Niech $s, t \geq 0$. Mamy

$$\mathbf{P}_x[T_1 > s + t] = \mathbf{E}_x[\mathbf{1}_{\{T_1 > s\}} \Phi \circ \theta_s],$$

gdzie $\Phi(\omega) = \mathbf{1}_{\{\omega(r) = \omega(0), \forall r \in [0, t]\}}$. Używając własności Markowa, dostajemy

$$\mathbf{P}_x[T_1 > s + t] = \mathbf{E}_x[\mathbf{1}_{\{T_1 > s\}} \mathbf{E}_x[\Phi]] = \mathbf{E}_x[\mathbf{1}_{\{T_1 > s\}} \mathbf{P}_x[T_1 > t]] = \mathbf{P}_x[T_1 > s] \mathbf{P}_x[T_1 > t],$$

co implikuje, że T_1 ma rozkład wykładniczy pod \mathbf{P}_x .

Założmy teraz, że $c(x) > 0$. Wówczas $T_1 < \infty$, \mathbf{P}_x prawie na pewno. Dla każdego $t \geq 0$ i $y \in S$,

$$\mathbf{P}_x[T_1 > t, X_{T_1} = y] = \mathbf{E}_x[\mathbf{1}_{\{T_1 > t\}} \Psi \circ \theta_t],$$

gdzie dla $\omega \in \Omega$, $\Psi(\omega) = 0$ jeśli ω jest stałe, a w przeciwnym razie $\Psi(\omega) = \mathbf{1}_{\{\gamma_1(\omega) = y\}}$, gdzie $\gamma_1(\omega)$ jest wartością ω po jego pierwszym skoku. Zatem mamy

$$\begin{aligned} \mathbf{P}_x[T_1 > t, X_{T_1} = y] &= \mathbf{E}_x[\mathbf{1}_{\{T_1 > t\}} \mathbf{E}_x[\Psi]] = \\ &= \mathbf{E}_x[\mathbf{1}_{\{T_1 > t\}} \mathbf{P}_x[X_{T_1} = y]] = \mathbf{P}_x[T_1 > t] \mathbf{P}_x[X_{T_1} = y], \end{aligned}$$

co daje pożądaną niezależność. \square

Punkty, dla których $c(x) = 0$, są stanami pochłaniającymi dla procesu Markowa, w tym sensie, że $\mathbf{P}_x[X_t = x, \forall t \geq 0] = 1$. Dla każdych $x, y \in S$ definiujemy

$$\Pi(x, y) = \begin{cases} \mathbf{P}_x[X_{T_1} = y] & c(x) > 0 \\ \delta_x(y) & c(x) = 0 \end{cases}$$

Zauważmy, że $\Pi(x, \cdot)$ jest miarą prawdopodobieństwa na S .

Twierdzenie 0.6. Niech (\mathbf{P}, \mathbb{F}) będzie łańcuchem Markowa w czasie ciągłym takim, że $\sup_{x \in S} c(x) < \infty$. Wówczas

$$\frac{d}{dt} \mathbb{P}_x[X_t = y]|_{t=0} = c(x)\Pi(x, y).$$

Proof. Jeśli $c(x) = 0$, to $\mathbf{P}_x[X_t = x] = \mathbf{P}_x[X_0 = x] = 1$, i stąd

$$\lim_{t \rightarrow 0} \frac{\mathbf{P}_x[X_t = x] - 1}{t} = 0.$$

Założmy teraz, że $c(x) > 0$. Najpierw zauważmy, że

$$\mathbf{P}_x[T_2 \leq t] = O(t^2) \quad (3)$$

gdy $t \rightarrow 0$. Rzeczywiście, używając silnej własności Markowa w T_1 ,

$$\mathbf{P}_x[T_2 \leq t] \leq \mathbf{P}_x[T_1 \leq t, T_2 \leq T_1 + t] = \mathbf{E}_x[\mathbf{1}_{\{T_1 \leq t\}} \mathbf{E}_{X_{T_1}}[T_1 \leq t]],$$

i możemy oszacować

$$\mathbf{P}_{X_{T_1}}[T_1 \leq t] \leq \sup_{y \in S} \mathbf{P}_y[T_1 \leq t] \leq \sup_{y \in S} c(y),$$

co daje oczekiwany wynik, ponieważ mamy również $\mathbf{P}_x[T_1 \leq t] \leq c(x)t$. Z (3) wynika, że

$$\begin{aligned} \mathbf{P}_x[X_t = y] &= \mathbf{P}_x[X_t = y, T_1 > t] + \mathbf{P}_x[X_{T_1} = y, T_1 \leq t] + O(t^2) \\ &= \delta_x(y)e^{-c(x)t} + (1 - e^{-c(x)t})\Pi(x, y) + O(t^2), \end{aligned}$$

używając niezależności T_1 i X_{T_1} oraz definicji $\Pi(x, y)$. Dochodzimy do wniosku, że skoro $\mathbf{P}_x[X_0 = y] = \delta_x(y)$, to

$$\frac{\mathbf{P}_x[X_t = y] - \mathbf{P}_x[X_0 = y]}{t} \rightarrow -c(x)\delta_x(y) + c(x)\Pi(x, y).$$

co kończy dowód. □

Kolejne twierdzenie dostarcza pełnego opisu ścieżek procesu X pod \mathbf{P}_x . Dla uproszczenia zakładamy, że nie ma stanów pochłaniających, ale czytelnik łatwo rozszerzy stwierdzenie na przypadek ogólny.

Twierdzenie 0.7. Zakładamy, że $c(y) > 0$ dla każdego $y \in S$ i że $\sup_{y \in S} c(y) < \infty$. Niech $x \in S$. Wówczas, \mathbf{P}_x p.n., czasy skoku $T_1 < T_2 < T_3 < \dots$ są skończone, a ciąg $X_0, X_{T_1}, X_{T_2}, \dots$ pod \mathbf{P}_x jest dyskretnym łańcuchem Markowa z macierzą przejścia Π rozpoczętym w x . Ponadto, pod warunkiem $(X_0, X_{T_1}, X_{T_2}, \dots)$, zmienne losowe $T_1 - T_0, T_2 - T_1, \dots$ są niezależne, a dla każdej liczby całkowitej $i \geq 0$, rozkład warunkowy $T_{i+1} - T_i$ jest wykładniczy z parametrem $c(X_{T_i})$.

Proof. Zastosowanie silnej własności Markowa pokazuje, że wszystkie czasy stopu T_1, T_2, \dots są skończone \mathbf{P}_x -p.n. Następnie, niech $y, z \in S$, a $f_1, f_2: S \rightarrow \mathbb{R}$. Używając silnej własności Markowa w T_1 :

$$\begin{aligned} \mathbf{E}_x[\mathbf{1}_{\{X_{T_1}=y\}} f_1(T_1) \mathbf{1}_{\{X_{T_2}=z\}} f_2(T_2 - T_1)] &= \mathbf{E}_x[\mathbf{1}_{\{X_{T_1}=y\}} f_1(T_1) \mathbf{E}_x[\mathbf{1}_{\{X_{T_2}=z\}} f_2(T_2 - T_1)]] \\ &= \Pi(x, y)\Pi(y, z) \int_0^\infty e^{-c(x)s_1} f_1(s_1) ds_1 \int_0^\infty e^{-c(y)s_2} f_2(s_2) ds_2. \end{aligned}$$

Postępując indukcyjnie, otrzymujemy dla każdych $y_1, \dots, y_p \in S$ oraz $f_1, \dots, f_p: S \rightarrow \mathbb{R}$:

$$\begin{aligned} \mathbf{E}_x[\mathbf{1}_{\{X_{T_1}=y_1\}} \mathbf{1}_{\{X_{T_2}=y_2\}} \dots \mathbf{1}_{\{X_{T_p}=y_p\}} f_1(T_1) f_2(T_2 - T_1) \dots f_p(T_p - T_{p-1})] \\ = \Pi(x, y_1) \Pi(y_1, y_2) \dots \Pi(y_{p-1}, y_p) \prod_{i=1}^p \left(\int_0^\infty e^{-c(y_{i-1})s} f_i(s) ds \right), \end{aligned}$$

gdzie $y_0 = x$. □

Z powyższego twierdzenia wynika charakteryzacja łańcucha Markowa w terminach Q -macierzy. Przez $\mathbb{F}^X = (\mathcal{F}_t^X)_t$ oznaczać będziemy najmniejszą możliwą filtrację, tj.

$$\mathcal{F}_t = \sigma(X_s : s \leq t).$$

Wniosek 0.1. *Niech $q = (q(x, y))_{x, y \in S}$ będzie Q -macierzą taką, że $\sup_{x \in S} |q(x, x)| < \infty$. Wówczas istnieje jedyna rodzina miar \mathbf{P} taka, że $(\mathbf{P}, \mathbb{F}^X)$ jest łańcuchem Markowa stowarzyszonym z Q -macierzą q .*

Wykład 3: procesy i półgrupy Fellera

2024-10-17

Piotr Dyszewski

W tym rozdziale S jest ośrodkową, lokalnie zwartą przestrzenią metryczną, a $C(S)$ jest przestrzenią ciągłych funkcji rzeczywistych na S . Przez $C_0(S)$ oznaczamy klasę funkcji z $C(S)$ znikających w nieskończoności. Dokładniej $C_0(S)$ to zbiór funkcji f z $C(S)$ takich, że dla każdego dodatniego ϵ istnieje zwarty $K \subseteq S$ taki, że $|f(x)| \leq \epsilon$ dla $x \in S \setminus K$. Zauważmy, że jeżeli S jest zwarta, to $C_0(S) = C(S)$. Dodatkowo każda f z $C_0(S)$ jest jednostajnie ciągła, tj. dla każdego dodatniego ϵ istnieje dodatnia δ , taka, że dla każdych $x, y \in S$,

$$d(x, y) < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \epsilon.$$

Tutaj d jest metryką na S . W obu przestrzeniach $C(S)$ i $C_0(S)$ używamy normy jednostajnej

$$\|f\| = \sup_{x \in S} |f(x)|,$$

co czyni $C_0(S)$ przestrzenią Banacha. Głównym powodem stosowania ciągłych funkcji znikających w nieskończoności zamiast ograniczonych ciągłych funkcji w przypadku lokalnie zwartym jest to, że jednostajna ciągłość jest wymagana w wielu argumentach. Ograniczone funkcje ciągłe nie są zwykle jednostajnie ciągłe, podczas gdy ciągłe funkcje znikające na nieskończoności są. Innym powodem jest to, że $C_0(S)$ jest ośrodkowa, co nie jest ogólnie prawdziwe dla przestrzeni wszystkich ograniczonych ciągłych funkcji na S .

Proces

Zaczynamy od opisu składników potrzebnych do definicji głównego obiektu zainteresowania w tym rozdziale. Konstrukcja będzie analogiczna do łańcuchów Markowa w czasie ciągłym. Niech $\Omega = D[0, \infty)$ będzie zbiorem funkcji prawostronnie ciągłych $\omega : [0, \infty) \rightarrow S$ z lewymi granicami w każdym punkcie. Tak jak poprzednio dla $s, t \in \mathbb{R}_+$ połóżmy też

$$X_t(\omega) = \omega(t) \text{ oraz } (\theta_s \omega)(t) = \omega(t + s).$$

Niech \mathcal{F} będzie najmniejszym σ -ciałem podzbiorów Ω względem którego wszystkie X_t dla $t \in \mathbb{R}_+$ są mierzalne.

Definicja 0.7. Procesem Fellera na S nazywamy parę uporządkowaną (\mathbf{P}, \mathbb{F}) taką, że

- **(PF1)** $\mathbf{P} = \{\mathbf{P}_x\}_{x \in S}$, gdzie dla każdego $x \in S$, \mathbf{P}_x jest miarą probabilistyczną na (Ω, \mathcal{F}) taką, że

$$\mathbf{P}_x[X_0 = x] = \mathbf{P}_x[\omega : \omega(0) = x] = 1. \quad (4)$$

- **(PF2)** $\mathbb{F} = \{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ jest filtracją na Ω , względem której zmienne losowe $X(t)$ są adaptowane.
- **(PF3)** Odwzorowanie

$$x \mapsto \mathbf{E}_x[f(X_t)] \text{ jest w } C_0(S) \text{ dla wszystkich } f \in C_0(S) \text{ i } t \geq 0. \quad (5)$$

- **(PF4)** Spełniona jest własność Markowa

$$\mathbf{E}_x[Y \circ \theta_s \mid \mathcal{F}_s] = \mathbf{E}_{X(s)}[Y] \quad \mathbf{P}_x\text{-prawie wszędzie} \quad (6)$$

dla wszystkich $x \in S$ oraz wszystkich ograniczonych mierzalnych Y na Ω .

Własność (5) znana jest jako własność Feller'a. Innym sposobem przedstawienia części ciągłości, który wydaje się całkiem naturalny, jest to, że $x_n \rightarrow x$ w S implikuje, że rozkład X dla procesu rozpoczynającego się w x_n zbiega się słabo do tego dla procesu rozpoczynającego się w x . Własność Feller'a (razem z prawostronną ciągłością trajektorii) implikuje silną własność Markowa.

Twierdzenie 0.8. *Każdy proces Feller'a ma silną własność Markowa. Jeżeli (\mathbf{P}, \mathbb{F}) jest procesem Feller'a, to dla każdej ograniczonej zmiennej $Y: \Omega \rightarrow \mathbb{R}$ oraz \mathbb{F} -czasu zatrzymania τ i każdego x ,*

$$\mathbf{E}_x[Y \circ \theta_\tau \mid \mathcal{F}_\tau] = \mathbf{E}_{X(\tau)}[Y] \quad \text{prawie na pewno } \mathbf{P}_x$$

na zdarzeniu $\{\tau < \infty\}$.

Zadanie 0.5. Niech (\mathbf{P}, \mathbb{F}) będzie procesem Feller'a. Pokaż, że odwzorowanie

$$x \mapsto \mathbf{E}_x \left[\prod_{j=1}^n f_j(X_{t_j}) \right] \quad (7)$$

jest ciągłe dla dowolnego n , dowolnych $t_1, \dots, t_n \in \mathbb{R}$ oraz dowolnych $f_1, \dots, f_n \in C_0(S)$.

Twierdzenia 0.8. Rozumowanie przebiega identycznie jak w przypadku łańcuchów Markowa w czasie ciągłym. W miejscu, w którym wymagana jest ciągłość odwzorowań $x \mapsto \mathbf{E}_x[Y]$ należy powołać się na tezę Zadania 0.5. \square

Przykład 0.3. Niech $B = (B_t)_{t \in \mathbb{R}_+}$ będzie standardowym ruchem Browna określonym na przestrzeni probabilistycznej $(\Sigma, \mathcal{G}, \mathbb{P})$. Przypomnijmy, że oznacza to, że

1. $B_0 = 0$ \mathbb{P} -p.w.
2. Dla dowolnych $t, s \in \mathbb{R}_+$, $t \geq s$ zmienna $B_t - B_s$ ma rozkład normalny $\mathcal{N}(0, t - s)$ o średniej zero i wariancji $t - s$.
3. Dla dowolnych $t, s \in \mathbb{R}_+$, $t \geq s$ zmienna $B_t - B_s$ jest niezależna od sigma ciała $\mathcal{G}_s^B = \sigma(B_r : r \leq s)$.
4. Odwzorowanie $t \mapsto B_t$ jest ciągłe.

Pokażemy, że ruch Browna jest procesem Feller'a w myśl przyjętej przez nas definicji. Połóżmy $\mathcal{F}_t = \sigma(X_s : s \leq t)$. Niech $S = \mathbb{R}$. Dla $x \in S$ zdefiniujemy \mathbf{P}_x jako rozkład ruchu Browna (rozumianego jako funkcji $\mathbb{R}_+ \rightarrow \mathbb{R}$) zapoczątkowanego w punkcie x , dokładniej dla $A \in \mathcal{F}$ niech $\mathbf{P}_x[A] = \mathbb{P}[B + x \in A]$. Tutaj przez $B + x$ rozumiemy funkcję $t \mapsto B_t + x$.

Wylosowanie ścieżki ω z rozkładu \mathbf{P}_x jest równoważne z wylosowaniem trajektorii ruchu Browna zapoczątkowanego w x .

Spełniona jest własność (PF1), ponieważ

$$\mathbf{P}_x[X_0 = x] = \mathbb{P}[B_0 + x = x] = 1.$$

Własność (PF2) jest spełniona wprost z definicji filtracji \mathbb{F} . Aby uzasadnić własność Feller'a (PF3) ustalmy $f \in C_0(S)$. Ciągłość

$$x \mapsto \mathbf{E}_x[f(X_t)] = \mathbb{E}[f(B_t + x)]$$

wynika z ciągłości f oraz twierdzenia o zbieżności ograniczonej. Aby uzasadnić, że powyższe odwzorowanie jest klasy $C_0(S)$ należy pokazać, że

$$\lim_{|x| \rightarrow \infty} \mathbf{E}_x[f(X_t)] = 0.$$

Wystarczy w tym celu rozważyć oszacowanie

$$|\mathbb{E}[f(B_t + x)]| \leq \|f\| \mathbb{P}[|B_t| > |x|/2] + \sup_{|y| > |x|/2} |f(y)|.$$

Oba składniki po prawej stronie zbiegają do zera, przy czym zbieżność tego drugiego wynika z $f \in C_0(\mathbb{R})$. Własność Markowa uzasadniamy dokładnie w taki sam sposób, w jaki zrobiliśmy to dla procesu Poissona w Przykładzie @ref{exm:2-poisson}.

Zadanie 0.6. Niech $S = \mathbb{Z}$. Pokaż, że łańcuch Markowa w czasie ciągłym (\mathbf{P}, \mathbb{F}) jest procesem Feller'a wtedy i tylko wtedy, gdy dla każdego $y \in S$ i każdego $t \in \mathbb{R}_+$,

$$\lim_{|x| \rightarrow \infty} \mathbf{P}_x[X_t = y] = 0.$$

Półgrupa

Chcemy teraz przedstawić odpowiednik funkcji przejścia na nieprzeliczalnej przestrzeni stanów. W naturalny sposób nasuwa się rozważenie rozkładów $\mathbf{P}_x[X_t \in dy]$. Jednak na dłuższą metę język rozkładów jest nieporęczny. O wiele bardziej praktyczny jest język półgrup. Aby umotywić następną definicję, rozważmy przeliczalną przestrzeń stanów S_0 oraz funkcję przejścia p na S_0 . Funkcję przejścia można zakodować w kategoriach rodziny operatorów

$$T_t f(x) = \sum_{y \in S_0} p_t(x, y) f(y), \quad (8)$$

dla $f \in C_0(S_0)$. Jasne jest, że znając T_t , a więc znając wartości $T_t f$ dla wszystkich $f \in C_0(S_0)$, znamy też funkcję przejścia $p_t(x, y)$. Wykorzystując równania Chapmana-Kołmogorowa dostajemy dla $s, t \geq 0$,

$$\begin{aligned} T_{s+t} f(x) &= \sum_{y \in S_0} p_{t+s}(x, y) f(y) = \sum_{y \in S_0} \sum_{z \in S_0} p_t(x, z) p_s(z, y) f(y) \\ &= \sum_{z \in S_0} p_t(x, z) \sum_{y \in S_0} p_s(z, y) f(y) = \sum_{z \in S_0} p_t(x, z) (T_s f)(z) = T_t(T_s f)(x). \end{aligned}$$

Wszystkie powyższe manipulacje są dozwolone ponieważ $f \in C_0(S_0)$ jest ograniczona. Powyższa tożsamość zapisuje się jako $T_t T_s = T_t \circ T_s = T_{t+s}$. Oznacza to, że $(T_t)_{t \geq 0}$ tworzą półgrupę.

Definicja 0.8. Półgrupa Feller'a to rodzina ciągłych operatorów liniowych $T = \{T_t\}_{t \in \mathbb{R}_+}$ na $C_0(S)$ spełniających następujące własności:

- a. $T_0 f = f$ dla wszystkich $f \in C_0(S)$.
- b. Dla każdego $f \in C_0(S)$, $\lim_{t \rightarrow 0} T_t f = f$ w $C_0(S)$.
- c. $T_{t+s} f = T_s T_t f$ dla każdego $f \in C_0(S)$.
- d. $T_t f \geq 0$ dla każdego nieujemnego $f \in C_0(S)$.
- e. Istnieje rodzina $f_n \in C_0(S)$, $n \in \mathbb{N}$ taka, że $\sup_n \|f_n\| < \infty$, oraz $T_t f_n$ zbiega punktowo do 1 dla każdego $t \geq 0$.

Część c. to analogia równań Chapmana-Kolmogorowa i nazywana jest własnością półgrupy. Jedną z jej konsekwencji jest to, że $T(t)$ i $T(s)$ komutują, tj. $T_t T_s = T_{t+s} = T_{s+t} = T_s T_t$. Z części d. i e. wynika, że $\|T(t)f\| \leq \|f\|$ dla wszystkich $f \in C_0(S)$, tak więc każdy $\|T\| \leq 1$. Własność b. jest znana jako mocna ciągłość. Wraz z c. i własnością kontrakcji, implikuje to, że funkcja $t \mapsto T(t)f$ z $[0, \infty)$ do $C_0(S)$ jest ciągła.

Oto ważny przykład - półgrupa Gaussa-Weierstrassa. Część b. tego ćwiczenia ilustruje powody przyjmowania funkcji w $C_0(S)$ zanikających na nieskończoności.

Zadanie 0.7. Niech $S = \mathbb{R}$ i $B = (B_t)_{t \in \mathbb{R}_+}$ będzie ruchem Browna.

- a. Pokaż, że T_t zdefiniowane przez

$$T_t f(x) = \mathbb{E}[f(B_t + x)]$$

jest półgrupą Feller'a.

- b. Wyjaśnij, dlaczego T nie jest mocno ciągła jako półgrupa operatorów na klasie $C_b(S)$ ograniczonych funkcji z $C(S)$.

W tym rozdziale konieczne będzie całkowanie funkcji ciągłych przyjmujących wartości w $C(S)$ względem t . Rachunek takich funkcji jest analogiczny do rachunku funkcji rzeczywistych. W tym duchu wiążemy z półgrupą jej transformata Laplace'a

$$U(\alpha)f = \int_0^\infty e^{-\alpha t} T_t f \, dt, \quad \alpha > 0, \quad (9)$$

która nazywana jest rezolwentą półgrupy. Funkcję $U(\alpha)f$ można interpretować jako całkę Bochnera pojawiającą się po prawej stronie (9). Można też równoważnie myśleć, że jest to funkcja $S \rightarrow \mathbb{R}$ zadana przez

$$U(\alpha)f(x) = \int_0^\infty e^{-\alpha t} T_t f(x) \, dt, \quad x \in S.$$

W każdym razie całka w (9) jest dobrze określona, ponieważ funkcja $t \mapsto e^{-\alpha t} T_t f$ jest ciągła oraz

$$\|e^{-\alpha t} T_t f\| \leq e^{-\alpha t} \|f\|.$$

Zauważmy też, że $U(\alpha)$ jest operatorem liniowym na $C_0(S)$ i spełnia

$$\|U(\alpha)f\| \leq \|f\|/\alpha.$$

Zadanie 0.8. Pokaż, że dla każdego $f \in C_0(S)$,

$$\lim_{\alpha \rightarrow \infty} \alpha U(\alpha) f = f.$$

Własność półgrupy przekłada się na następującą użyteczną relację, znaną jako równanie rezolwenty:

$$U(\alpha) - U(\beta) = (\beta - \alpha)U(\alpha)U(\beta). \quad (10)$$

Aby to sprawdzić, weźmy $\alpha \neq \beta$ i zapiszmy

$$\begin{aligned} U(\alpha)U(\beta)f &= \int_0^\infty e^{-\alpha t} T_t U(\beta)f \, dt = \int_0^\infty e^{-\alpha t} \left(\int_0^\infty e^{-\beta s} T_t T_s f \, ds \right) dt \\ &= \int_0^\infty \int_0^r e^{-\alpha t} e^{-\beta(r-t)} \, dt T_r f \, dr = \int_0^\infty \frac{e^{-\alpha r} - e^{-\beta r}}{\beta - \alpha} T_r f \, dr. \end{aligned} \quad (11)$$

Jedną z konsekwencji (10) jest to, że $U(\alpha)$ i $U(\beta)$ komutują.

Wykład 4: Generatory

2024-10-24

Piotr Dyszewski

Do tej pory te definicje powinny być dość intuicyjne. Kolejna definicja może wydawać się mniej oczywista, jednak okazuje się być odpowiednim odpowiednikiem definicji macierzy Q .

Aby nakreślić analogię, biorąc macierz Q na przeliczalnym zbiorze S_0 , niech p będzie funkcją przejścia zadaną jako

$$p_t = e^{tq} = \sum_{k=0}^{\infty} \frac{t^k}{k!} q^k.$$

Z tą funkcją wiążemy półgrupę

$$T_t f(x) = \sum_{y \in S_0} p_t(x, y) f(y),$$

a także rezolwentę

$$U(\alpha) f(x) = \int_0^{\infty} e^{-\alpha t} T_t f(x) dt = \int_0^{\infty} e^{-\alpha t} e^{tq} f(x) dt = \int_0^{\infty} e^{-\alpha t} e^{tq} dt f(x).$$

Ostatnie przejście wynika z faktu, że mamy tutaj do czynienia z mnożeniem wektora przez macierz. Zauważmy, że

$$(q - \alpha I) \int_0^{\infty} e^{-\alpha t} e^{tq} dt = \int_0^{\infty} (q - \alpha I) e^{-\alpha t} e^{tq} dt = \int_0^{\infty} \frac{d}{dt} e^{-\alpha t} e^{tq} dt = -I.$$

Oznacza to, że

$$U(\alpha) f(x) = (\alpha I - q)^{-1} f(x).$$

Z własności rezolwenty wiemy, że

$$\left\| \left(I - \frac{1}{\alpha} q \right)^{-1} \right\| = \|\alpha U(\alpha)\| \leq 1.$$

Ostatnia własność rezolwenty, z której tutaj skorzystaliśmy, wynika z kontraktywności operatorów w półgrupie T ($\|T_t\| \leq 1$).

Definicja 0.9. Generator infinitesimalny na $C_0(S)$ to para uporządkowana $(L, \mathcal{D}(L))$ taka, że:

- **(GI1)** $\mathcal{D}(L)$ jest gęstą podprzestrzenią liniową $C_0(S)$.
- **(GI2)** $L: \mathcal{D}(L) \rightarrow C_0(S)$ jest operatorem liniowym.
- **(GI3)** Jeśli $f \in \mathcal{D}(L)$, $\lambda \geq 0$, i $f - \lambda Lf = g$, to

$$\inf_{x \in S} f(x) \geq \inf_{x \in S} g(x).$$

- **(GI4)** $\mathcal{R}(I - \lambda L) = C_0(S)$ dla wszystkich dostatecznie małych $\lambda > 0$.
- **(GI5)** Dla dostatecznie małych dodatnich λ istnieje ciąg $f_n \in \mathcal{D}(L)$ (który może zależeć od λ) taki, że $g_n = f_n - \lambda L f_n$ spełnia warunek $\sup_n \|g_n\| < \infty$, i zarówno f_n , jak i g_n zbiega punktowo do 1.

Zauważmy, że własność **(GI3)** ma następującą konsekwencję:

$$f \in \mathcal{D}(L), \lambda \geq 0, f - \lambda L f = g \implies \|f\| \leq \|g\|. \quad (12)$$

Aby to zobaczyć, napiszmy:

$$\inf_{x \in S} g(x) \leq \inf_{x \in S} f(x) \leq \sup_{x \in S} f(x) \leq \sup_{x \in S} g(x),$$

gdzie ostatnia nierówność wynika z **(GI3)**, gdy zastąpimy f i g odpowiednio przez $-f$ i $-g$. Oznacza to, że operator $I - \lambda L$ jest różnowartościowy. Rzeczywiście, dla $f - \lambda L f = g = h - \lambda L h$, mamy $\|f - h\| \leq \|g - g\| = 0$. Tak więc, dla dostatecznie małych dodatnich λ , $(I - \lambda L)^{-1}$ jest dobrze określoną kontrakcją, która odwzorowuje funkcje nieujemne na funkcje nieujemne.

Ponieważ Definicja @ref{def:3-12} jest dość abstrakcyjna, pomocne może być rozważenie następującego przykładu, który okazuje się być generatorem procesu na prostej, poruszającego się w prawo z jednostkową prędkością. Zauważmy, że najtrudniejszą własnością do sprawdzenia jest **(GI4)**. Zazwyczaj tak bywa.

Zadanie 0.9. Przypuśćmy, że $S = \mathbb{R}$,

$$\mathcal{D}(L) = \{f \in C_0(\mathbb{R}) : f' \in C_0(\mathbb{R})\},$$

oraz $Lf = f'$. Pokaż, że para $(L, \mathcal{D}(L))$ jest generatorem infinitezimalnym.

Od procesu do półgrupy i generatora

Oto pierwszy krok w przejściu od procesu Fellera do jego generatora.

Twierdzenie 0.9. Niech dany będzie proces Fellera (\mathbf{P}, \mathbb{F}) . Dla $t \geq 0$ zdefiniujmy

$$T_t f(x) = \mathbf{E}_x[f(X(t))] \quad (13)$$

dla $f \in C_0(S)$. Wtedy $T = (T_t)_{t \geq 0}$ jest półgrupą Fellera.

Proof. Własności a., d. i e. z Definicji 0.8 są natychmiastowe. Własność półgrupy c. wynika z własności Markowa:

$$T_{s+t} f(x) = \mathbf{E}_x f(X(s+t)) = \mathbf{E}_x[\mathbf{E}_{X(s)} f(X(t)) | \mathcal{F}_s] = \mathbf{E}_x[T_t f(X(s))] = T_s T_t f(x).$$

Zbieżność punktowa w b. wynika z ciągłości ścieżek i ciągłości f . Aby sprawdzić wymaganą jednostajność w tej zbieżności, użyjemy rezolwenty

$$U(\alpha) f(x) = \int_0^\infty e^{-\alpha t} T_t f(x) dt = \mathbf{E}_x \left[\int_0^\infty e^{-\alpha t} f(X_t) dt \right].$$

W dowodzie równania rezolwenty (10), użyliśmy wspomnianej jednostajności, ponieważ całki były interpretowane jako całki funkcji o wartościach w $C_0(S)$. Jednak te same obliczenia stosują się

do równania rezolwenty bez tej jednostajności, jeśli całki są interpretowane jako zwykłe całki dla ustalonego x . Aby uzasadnić zamianę kolejności całkowania, zauważmy, że $T_t f(x)$ jest jednostajnie ograniczone, prawostronnie ciągłe w t dla każdego x , a także ciągłe w x dla każdego t , zatem jest wspólnie mierzalne względem x i t .

Zbiór $\mathcal{L} = \mathcal{R}(U(\alpha))$ jest niezależny od α . Z równania rezolwenty mamy bowiem

$$U(\alpha)f = U(\beta)(f + (\beta - \alpha)U(\beta)f).$$

Jeśli $f = U(\alpha)g \in \mathcal{L}$, to

$$T_t f = \int_0^\infty e^{-\alpha s} T_{s+t} g \, ds = \int_t^\infty e^{-\alpha(r-t)} T_r g \, dr,$$

co zbiega jednostajnie do f gdy $t \downarrow 0$. Ponieważ każde T_t jest kontrakcją, mamy $\|T_t f - f\| \rightarrow 0$ dla wszystkich f w mocnym domknięciu $\text{cl}(\mathcal{L})$. Pokażemy teraz, że wspomniane mocne domknięcie jest równe słabemu domknięciu:

$$\text{wcl}(\mathcal{L}) = \{f \in C_0(S) : \exists \{f_n\}_n \subseteq \mathcal{L}, x(f_n) \rightarrow x(f), \forall x \in C_0(S)^*\}.$$

Skoro zbieżność w normie implikuje zbieżność słabą, to $\text{cl}(\mathcal{L}) \subseteq \text{wcl}(\mathcal{L})$. Weźmy $f \notin \text{cl}(\mathcal{L})$. Ponieważ \mathcal{L} jest wypukły, z twierdzenia Hahna-Banacha istnieje funkcjonal liniowy μ oddzielający f od $\text{cl}(\mathcal{L})$, taki że

$$\mu(f) < \inf_{g \in \text{cl}(\mathcal{L})} \mu(g).$$

Funkcjonał ten dowodzi, że $f \notin \text{wcl}(\mathcal{L})$.

Pozostaje pokazać, że słabe domknięcie \mathcal{L} jest równe całemu $C_0(S)$. Jest tak, ponieważ $\alpha U(\alpha)f$ zbiega punktowo do f gdy $\alpha \rightarrow \infty$ dla każdego $f \in C_0(S)$. \square

Następnie zobaczymy, jak przejść od półgrupy do generatora.

Twierdzenie 0.10. *Przypuśćmy, że $(T_t)_{t \geq 0}$ jest półgrupą Fellera. Zdefiniujmy*

$$Lf = \lim_{t \downarrow 0} \frac{T(t)f - f}{t} \quad (14)$$

dla f z

$$\mathcal{D}(L) = \{f \in C(S) : \text{przy } t \rightarrow 0 \text{ granica } (T_t f - f)/t \text{ istnieje}\}.$$

Wtedy para $(L, (\mathcal{D}(L)))$ jest generatorem infinitezymalnym. Ponadto:

a. Dla dowolnego $g \in C_0(S)$ oraz $\alpha > 0$,

$$f = \alpha U(\alpha)g \text{ wtedy i tylko wtedy, gdy } f \in \mathcal{D}(L) \text{ i spełnia } f - \alpha^{-1}Lf = g. \quad (15)$$

b. Jeśli $f \in \mathcal{D}(L)$, to $T_t f \in \mathcal{D}(L)$ dla wszystkich $t \geq 0$, jest funkcją ciągłą, różniczkowalną względem t , i spełnia

$$\frac{d}{dt} T_t f = T_t Lf = L T_t f. \quad (16)$$

Proof. Przypuśćmy, że $f = \alpha U(\alpha)g$ dla pewnego $\alpha > 0$ oraz $g \in C_0(S)$. Korzystając z własności półgrupy i zmieniając zmienne jak w dowodzie Twierdzenia @ref{thm:3-15}, mamy:

$$\begin{aligned} \frac{T_t f - f}{t} &= \alpha \frac{e^{\alpha t} - 1}{t} \int_t^\infty e^{-\alpha s} T_s g \, ds - \alpha \frac{1}{t} \int_0^t e^{-\alpha s} T_s g \, ds \\ &\rightarrow \alpha^2 U(\alpha)g - \alpha g = \alpha f - \alpha g, \end{aligned}$$

gdy $t \downarrow 0$. Przy przejściu do granicy skorzystaliśmy z własności b. z Definicji @ref{def:3-1}. To dowodzi jednej implikacji w (15), jak również **(GI4)** w Definicji 0.9.

Ponieważ $\alpha U(\alpha)g \in \mathcal{D}(L)$ i $\alpha U(\alpha)g \rightarrow g$ gdy $\alpha \rightarrow \infty$, zbiór $\mathcal{D}(L)$ jest gęsty w $C_0(S)$. To uzasadnia **(GI1)**.

Dla $t > 0$ oraz $f \in \mathcal{D}(L)$ zdefiniujemy

$$g_t = \left(1 + \frac{\lambda}{t}\right) f - \frac{\lambda}{t} T(t)f = f - \frac{\lambda}{t} (T(t)f - f).$$

Wtedy $\lim_{t \downarrow 0} g_t = f - \lambda \mathcal{L}f$ i

$$(1 + \lambda/t) \inf_{x \in S} f(x) \geq \frac{\lambda}{t} \inf_{x \in S} T(t)f(x) + \inf_{x \in S} g_t(x) \geq \frac{\lambda}{t} \inf_{x \in S} f(x) + \inf_{x \in S} g_t(x),$$

więc własność **(GI3)** z Definicji 0.9 jest spełniona. Drugą nierówność pozostawiamy jako zadanie.

Teraz przypuśćmy, że $f - \alpha^{-1} Lf = g$ dla pewnego $f \in \mathcal{D}(L)$ oraz $\alpha > 0$. Przez dowiedzioną już implikację w (15), $h = \alpha U(\alpha)g$ spełnia $h - \alpha^{-1} Lh = g$, więc $f = h$ z (12). Aby sprawdzić własność **(GI5)** z Definicji 0.9, przypuśćmy, że $g_n \in C(S)$ spełniają $\sup_n \|g_n\| < \infty$, oraz że g_n i $T(t)g_n$ są zbieżne punktowo do 1 dla każdego t . Zdefiniujemy $f_n \in \mathcal{D}(L)$ przez $g_n = f_n - \lambda Lf_n$. Wtedy $f_n = \alpha U(\alpha)g_n$ z (15). Ponieważ $T(t)g_n \rightarrow 1$ punktowo, to $f_n \rightarrow 1$ punktowo przez definicję $U(\alpha)$ oraz twierdzenie o zbieżności ograniczonej.

Aby udowodnić punkt (b) twierdzenia, zauważmy, że

$$\frac{d}{dt} T(t)f = \lim_{s \rightarrow 0} \frac{T(t+s)f - T(t)f}{s} = \lim_{s \rightarrow 0} T(t) \left(\frac{T(s)f - f}{s} \right) = T(t) \mathcal{L}f = \mathcal{L}T(t)f,$$

pod warunkiem, że którakolwiek z granic istnieje, ponieważ wyrażenia w środku granic są identyczne. Środkowa granica rzeczywiście istnieje, ponieważ $f \in \mathcal{D}(L)$ oraz $T(t)$ jest kontrakcją. W związku z tym pozostałe granice również istnieją.

Skoro istnieje trzecia granica, to $T(t)f \in \mathcal{D}(L)$ oraz (16) zachodzi. Środkowe wyrażenie w (16) jest ciągle względem t , więc $T(t)f$ jest ciągle i różniczkowalne. \square

Wykład 5: od generatora do półgrupy

2024-10-31

Piotr Dyszewski

O notacji słów kilka

Twierdzenie 0.11. Przy oznaczeniach Twierdzenia @0.10, dla $f \in C_0(S)$ oraz $t > 0$,

$$\lim_{n \rightarrow \infty} \left(I - \frac{t}{n} L \right)^{-n} f = T_t f.$$

Proof. Sprawdzamy indukcyjnie, że

$$\left(I - \alpha^{-1} L \right)^{-n} f = \alpha^n U^n(\alpha) f = \int_0^\infty \alpha^n s^{n-1} e^{-\alpha s} T(s) f \, ds.$$

Stąd

$$\left(I - \frac{t}{n} L \right)^{-n} f = \mathbb{E} T((\xi_1 + \dots + \xi_n)t/n) f, \quad (17)$$

gdzie ξ_1, ξ_2, \dots są niezależnymi zmiennymi o standardowym rozkładzie wykładniczym. Zauważmy, że funkcja $(s, x) \rightarrow T(s)f(x)$ jest ciągła, więc wartość oczekiwana w (17) jest dobrze określona.

Jeśli $f \in \mathcal{D}(L)$, to

$$\frac{d}{dt} T_t f = T_t L f$$

co w notacji całkowej zapisuje się jako

$$T_t f - T_s f = \int_s^t T_r L f \, dr.$$

Skoro $\|T_r\| \leq 1$, to ostatnia nierówność implikuje, że

$$\|T(t)f - T(s)f\| \leq \|Lf\| |t - s|.$$

Wracając teraz do (17) otrzymujemy

$$\left\| \left(I - \frac{t}{n} L \right)^{-n} f - T(t)f \right\| \leq t \|Lf\| \mathbb{E} \left| \frac{\xi_1 + \dots + \xi_n}{n} - 1 \right|.$$

Rezultat dla $f \in \mathcal{D}(L)$ wynika teraz z prawa wielkich liczb. Jest on prawdziwy dla wszystkich $f \in C_0(S)$, ponieważ wszystkie rozważane operatory są kontrakcjami. \square

Remark. Formalnie, $T_t = \exp(tL)$. Kiedy L jest ograniczone, istnieją przynajmniej trzy sposoby definiowania tego wykładnika:

$$\sum_{k=0}^{\infty} \frac{(tL)^k}{k!}, \quad \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} L \right)^n, \quad \text{ i } \quad \lim_{n \rightarrow \infty} \left(I - \frac{t}{n} L \right)^{-n}.$$

Ostatni z nich jest jedynym, który ma sens w przypadku nieograniczonym.

Teraz rozważmy kilka przykładów.

Zadanie 0.10. Rozważmy proces Feller'a polegający na jednostajnym ruchu w prawo. Niech $S = \mathbb{R}$ i niech \mathbf{P}_x jest punktową masą na ścieżce ω_x danej przez $\omega_x(t) = x + t$, lub równoważnie, proces z półgrupą $T(t)f(x) = f(x+t)$. Pokaż, że generatorem L tego procesu jest ten opisany w Zadaniu~??. Upewnij się, że dziedzina dana jest dokładnie przez dziedzinę L .

Przykład 0.4. W przypadku ruchu Browna, można zweryfikować, że

$$U(\lambda)f(x) = \int u_\lambda(y-x)f(y) dy$$

gdzie

$$u_\lambda(y-x) = \int_0^\infty (2\pi t)^{-1/2} \exp\left(-\frac{|y-x|^2}{2t} - \lambda t\right) dt = \frac{1}{\sqrt{2\lambda}} \exp(-|y-x|\sqrt{2\lambda}).$$

Elegancki sposób na uzyskanie tej ostatniej równości polega na użyciu wzoru $\mathbb{E}[e^{-\lambda T_a}] = e^{-a\sqrt{2\lambda}}$ dla transformaty Laplace'a czasu trafienia $a > 0$ przez rzeczywisty ruch Browna rozpoczęty z 0. Różniczkując względem λ , otrzymujemy

$$\mathbb{E}[T_a e^{-\lambda T_a}] = \frac{a}{\sqrt{2\lambda}} e^{-a\sqrt{2\lambda}},$$

i używając gęstości T_a , aby przekształcić $\mathbb{E}[T_a e^{-\lambda T_a}]$, dokładnie znajdujemy całkę, która pojawia się w obliczeniach $u_\lambda(y-x)$.

Wiemy, że półgrupa operatorów związana z ruchem Browna jest półgrupą Feller'a. Znajdziemy jej generator L . Widzieliśmy, że dla każdego $\lambda > 0$ i $f \in C_0(\mathbb{R})$,

$$U(\lambda)f(x) = \int \frac{1}{\sqrt{2\lambda}} \exp(-\sqrt{2\lambda}|y-x|)f(y) dy.$$

Jeśli $h \in \mathcal{D}(L)$, wiemy, że istnieje $f \in C_0(\mathbb{R})$ takie, że $h = U(\lambda)f$. Przyjmując $\lambda = 1/2$, mamy

$$h(x) = \int \exp(-|y-x|)f(y) dy.$$

Różniczkując pod znakiem całki (pozostawiamy uzasadnienie jako zadanie), otrzymujemy, że h jest różniczkowalna na \mathbb{R} , i

$$h'(x) = \int \operatorname{sgn}(y-x) \exp(-|y-x|)f(y) dy$$

gdzie $\operatorname{sgn}(z) = 1_{\{z>0\}} - 1_{\{z<0\}}$ (wartość $\operatorname{sgn}(0)$ jest nieistotna). Pokażmy również, że h' jest różniczkowalna na \mathbb{R} . Niech $x_0 \in \mathbb{R}$. Następnie, dla $x > x_0$,

$$\begin{aligned} h'(x) - h'(x_0) &= \int_{\mathbb{R}} (\operatorname{sgn}(y-x) \exp(-|y-x|) - \operatorname{sgn}(y-x_0) \exp(-|y-x_0|)) f(y) dy \\ &= \int_{x_0}^x (-\exp(-|y-x|) - \exp(-|y-x_0|)) f(y) dy \\ &\quad + \int_{\mathbb{R} \setminus [x_0, x]} \operatorname{sgn}(y-x_0) (\exp(-|y-x|) - \exp(-|y-x_0|)) f(y) dy. \end{aligned}$$

Wynika stąd, że

$$\frac{h'(x) - h'(x_0)}{x - x_0} \xrightarrow{x \downarrow x_0} -2f(x_0) + h(x_0).$$

Otrzymujemy tę samą granicę, gdy $x \uparrow x_0$, i stąd uzyskujemy, że h jest dwukrotnie różniczkowalna oraz $h'' = -2f + h$.

Z drugiej strony, ponieważ $h = U(1/2)f$,

$$\left(\frac{1}{2} - L\right)h = f$$

stąd $Lh = -f + h/2 = h''/2$. Podsumowując, uzyskaliśmy, że

$$\mathcal{D}(L) \subset \{h \in C^2(\mathbb{R}) : h \text{ i } h'' \in C_0(\mathbb{R})\}$$

i że, jeśli $h \in \mathcal{D}(L)$, mamy $Lh = h''/2$.

Ostatnie zawieranie jest w rzeczywistości równością. Aby to zobaczyć, możemy postąpić następująco. Jeśli g jest funkcją dwukrotnie różniczkowalną taką, że g oraz g'' należą do $C_0(\mathbb{R})$, wtedy przyjmujemy $f = \frac{1}{2}(g - g'') \in C_0(\mathbb{R})$, a więc $h = U(1/2)f \in \mathcal{D}(L)$. Z poprzedniego argumentu wynika, że h jest dwukrotnie różniczkowalna oraz $h'' = -2f + h$. Stąd $(h - g)'' = h - g$. Ponieważ funkcja $h - g$ należy do $C_0(\mathbb{R})$, musi zanikać tożsamościowo, co daje $g = h \in \mathcal{D}(L)$.

Patrząc na Zadanie 0.9 i Przykład 0.4, można by się zastanawiać, czy pochodne wyższego rzędu mogą być generatorami infinitezymalnymi. Odpowiedź brzmi, że nie mogą. Główny problem polega na tym, że gdy gładka funkcja osiąga minimum w punkcie wewnętrznym swojej dziedziny, pierwsza pochodna jest zerowa, a druga pochodna jest tam nieujemna. Nic nie można powiedzieć o znakach innych pochodnych w tym miejscu.

Zadanie 0.11. Pokaż, że nie istnieje generator prawdopodobieństwa, którego ograniczenie do gładkich funkcji jest dane przez $Lf = f'''$.

Od generatora do procesu

W całym tym rozdziale L będzie generatorem infinitezymalnym. Naszym pierwszym zadaniem jest skonstruowanie odpowiadającej półgrupy prawdopodobieństwa. Aby to zrobić, wprowadzamy aproksymację L_ϵ do L dla małego dodatniego ϵ przez

$$L_\epsilon = L(I - \epsilon L)^{-1}.$$

Zauważmy, że jest to dobrze zdefiniowane z definicji generatora infinitezymalnego, ponieważ $\mathcal{R}(I - \epsilon L) = \mathcal{D}(L)$. To łatwo zobaczyć z

$$f - \epsilon Lf = g \quad \text{jest równoważne} \quad f = (I - \epsilon L)^{-1}g.$$

Ponadto

$$\|L_\epsilon g\| = \|Lf\| \leq \frac{\|f\| + \|g\|}{\epsilon} \leq \frac{2}{\epsilon} \|g\|,$$

więc L_ϵ jest ograniczonym operatorem. To pozwala zdefiniować $T_\epsilon(t)$ przez

$$T_\epsilon(t) = e^{tL_\epsilon} = \sum_{n=0}^{\infty} \frac{t^n L_\epsilon^n}{n!}.$$

Zadanie 0.12.

- (a) Pokaż, że dla każdego $f \in C_0(S)$,

$$(I - \epsilon L)^{-1} f - \epsilon L f = f. \quad (18)$$

- (b) Użyj części (a), aby pokazać, że L_ϵ jest generatorem infinitesimalnym oraz że $T_\epsilon(t)$ jest półgrupą prawdopodobieństwa, której generatorem jest L_ϵ w sensie Twierdzenia 0.10.

Twierdzenie 0.12. Dla $f \in C_0(S)$,

$$T(t)f = \lim_{\epsilon \rightarrow 0} T_\epsilon(t)f$$

jednostajnie na ograniczonych przedziałach t . Definiuje to półgrupę Fellera, której generatorem jest L w sensie Twierdzenia 0.10.

Najpierw sprawdzamy, że L_ϵ i L_δ komutują dla $\epsilon, \delta > 0$. Wynika to z @ref{eq:3-19} oraz

$$(I - \epsilon L)^{-1}(I - \delta L)^{-1} = (I - \delta L)^{-1}(I - \epsilon L)^{-1},$$

co jest prawdziwe, ponieważ

$$(I - \epsilon L)^{-1}(I - \delta L)^{-1}f = g \quad \text{jest równoważne} \quad f = g - (\epsilon + \delta)Lg + \epsilon\delta L^2g,$$

co jest symetryczne w ϵ i δ .