

Stochastyczne modele układów oddziałujących 2024

Notatni do wykładu

Piotr Dyszewski

2024-12-27

Contents

Stochastyczne modele układów oddziałujących	2
Sylabus	3
Skrócony plan wykładu	3
Szczegółowy plan wykładu	4
Wykład 1: Łańcuchy Markowa w czasie ciągłym	6
Podstawowe definicje	8
Wykład 2: Mocna własność Markowa	19
Czasy zatrzymania	19
Charakteryzacja	23
Symulacja łańcucha Markowa w czasie ciągłym	26
Wykład 3: procesy i półgrupy Feller'a	29
Proces	30
Półgrupa	33
Wykład 4: Generatory	36
Od procesu do półgrupy i generatora	38
Wykład 5: od generatora do półgrupy	42
O notacji słów kilka	42
Od generatora do procesu	45
Wykład 6: od generatora przez półgrupę do procesu	47
Wykład 7: Rozkłady stacjonarne, zaburzenia ruchu Browna	54

Rozkłady stacjonarne	54
Zaburzenia ruchu Browna	57
Wykład 8: Układy spinowe	61
Motywacje	61
Systemy spinowe	63
Dygresja analityczna	63
Konstrukcja systemów spinowych	66
wykład 9: Konstrukcja systemów spinowych	67
Konstrukcja systemów spinowych cd.	67
Wykład 10: Procesy dualne	72
Procesy dualne	72
The Voter model	75
Wykład 11: The voter model: przypadek rekurencyjny	79
The Voter model na \mathbb{Z}^d	80
Przypadek rekurencyjny	80

Stochastyczne modele układów oddziałujących

Notatki te powstały na potrzebę kursu o tym samym tytule prowadzonym w Instytucie Matematycznym Uniwersytetu Wrocławskiego zimą 2024. Zainspirowany *Complex analysis* autorstwa Juan Carlos Ponce Campuzano oraz *Collision Detection* Jeffreya Thompsona chciałem zbadać możliwości tworzenia interaktywnych podręczników. Wierzę, że wchodzenie w interakcję z czytelnikiem pozwala mu lepiej zrozumieć prezentowany materiał. Jednym z najlepszych sposobów na zrozumienie nawet najbardziej zawiłych procesów jest zobaczenie ich w akcji.

Notatki te w dalszym ciągu są we wstępnej fazie i niestety nie są wolne od błędów. Gorąco zachęcam do zgłaszania wszystkich problemów, błędów i niedociągnięć na [github](#).

Do napisania książki wykorzystałem bibliotekę `bookdown` z pakietu R. Symulacje zostały napisane z wykorzystaniem `p5.js`.

Wrocław, grudzień 2024
Piotr Dyszewski

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 Feller'a, półgrupy operatorów, generatory infinitesimalne, 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 Feller'a, 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 infinitezymalny. 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”. cld

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

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

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

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

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

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

Relacja powyższa daje się zapisać w przypadku nieprzeliczalnej przestrzeni stanów jako

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

gdzie

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

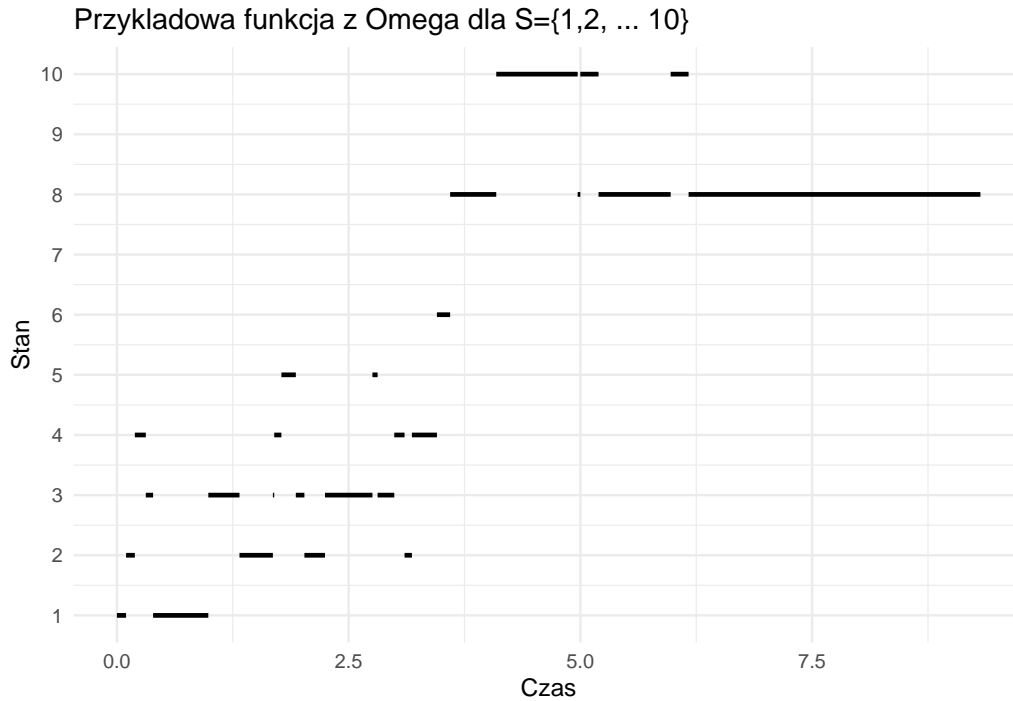
jest rozkładem X_t pod warunkiem $\{X_0 = y\}$. To niesie ze sobą kolejne problemy, ponieważ jak wcześniej zauważyliśmy $\{X_0 = y\}$ może być zdarzeniem o prawdopodobieństwie zero. Przedstawione podejście jest do uratowania pod kątem formalnym przez odniesienie się do regularnych rozkładów warunkowych.

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.



```
## List of 1
## $ plot.margin: 'margin' num [1:4] 0mm 25mm 0mm 0mm
## ..- attr(*, "unit")= int 7
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi FALSE
## - attr(*, "validate")= logi TRUE
```

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 mierzalne dla każdego $t \in \mathbb{R}_+ = [0, +\infty)$. Niech wreszcie dla $s \in \mathbb{R}_+$ oznaczymy 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

$$\begin{aligned} Y \circ \theta_s(\omega) &= f(\omega(t_1 + s), \omega(t_2 + s), \dots, \omega(t_n + s)) \\ &= f(X(t_1 + s), X(t_2 + s), \dots, X(t_n + s)). \end{aligned}$$

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 $\mathcal{F}_s^N = \sigma(N_r : r \leq s)$.

Parametr lambda w procesie Poissona jest odpowiedzialny za tempo jego wzrostu. Dla małych wartości parametru lambda, odstęp między skokami są duże i proces porusza się wolno. Dla dużych wartości parametru lambda odstęp są mniejsze i proces porusza się szybciej. Za pomocą poniższego suwaka możesz ustawić wartość parametru lambda w zakresie od 0.1 do 1.

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-Kolmogorowa 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

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

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. *Załóżmy, 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

$$\begin{aligned} \sum_{z \notin T} \left| \frac{p_s(x, z)}{s} - q(x, z) \right| p_t(z, y) &\leq \sum_{z \notin T} \frac{p_s(x, z)}{s} + \sum_{z \notin T} q(x, z) \\ &= s^{-1} \left[1 - \sum_{z \in T} p_s(x, z) \right] - \sum_{z \in T} q(x, z) \rightarrow -2 \sum_{z \in T} q(x, z) \end{aligned}$$

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 τ , τ_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:

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

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,

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

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

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

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

Załóżmy 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*

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

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

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

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. \square

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 .*

Symulacja łańcucha Markowa w czasie ciągłym

W tej sekcji przedstawimy, jak zaimplementować i zwizualizować łańcuch Markowa w czasie ciągłym z 5 stanami.

```
# Funkcja do symulacji łańcucha Markowa w czasie ciągłym
simulate_markov_chain <- function(Q, initial_state, time_horizon) {
  states <- nrow(Q)
  state <- initial_state
  time <- 0
  times <- c(0)
  trajectory <- c(state)

  while (time < time_horizon) {
    rate <- -Q[state, state] # Intensywność opuszczenia stanu
    if (rate == 0) {
      break
    }
    waiting_time <- rexp(1, rate) # Czas do następnego zdarzenia
    time <- time + waiting_time
    if (time > time_horizon) {
      break
    }
    times <- c(times, time)

    # Prawdopodobieństwa przejść do innych stanów
    probs <- Q[state, ]
```

```

    probs[state] <- 0 # Ustawiamy intensywność własną
    probs <- probs / sum(probs)

    state <- sample(1:states, 1, prob = probs) # Nowy stan
    trajectory <- c(trajectory, state)
  }

  return(data.frame(time = times, state = trajectory))
}

# Macierz przejść (generator Q)
Q <- matrix(c(-2, 1, 1, 0, 0,
              1, -3, 1, 1, 0,
              1, 1, -3, 0, 1,
              0, 1, 0, -2, 1,
              0, 0, 1, 1, -2), byrow = TRUE, nrow = 5)

# Symulacja łańcucha Markowa
set.seed(42) # Ustalenie ziarna dla powtarzalności wyników
initial_state <- 1
czas_horyzont <- 10
data <- simulate_markov_chain(Q, initial_state, czas_horyzont)

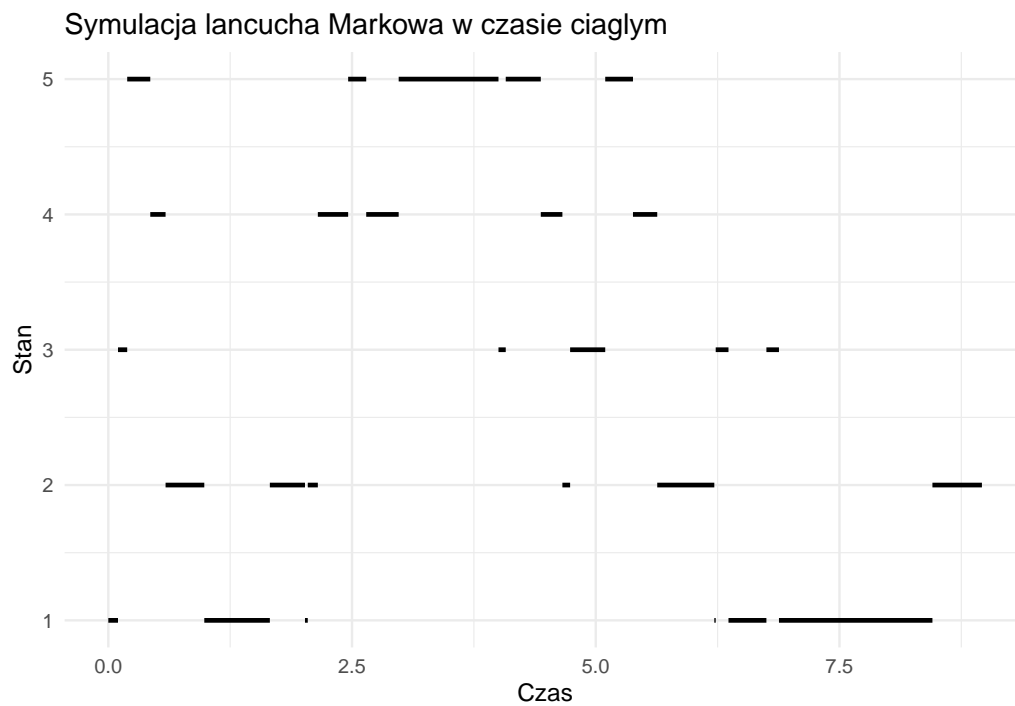
# Przygotowanie danych do wykresu kawałkami stałego
prepare_data_for_step <- function(data) {
  result <- data.frame(
    x_start = head(data$time, -1),
    x_end = tail(data$time, -1),
    state = head(data$state, -1)
  )
  return(result)
}

# Przygotowanie danych
step_data <- prepare_data_for_step(data)

# Tworzenie wykresu kawałkami stałego

```

```
ggplot(step_data, aes(x = x_start, y = state)) +
  geom_segment(aes(xend = x_end, yend = state), size = 1) +
  scale_y_continuous(breaks = 1:5) +
  labs(
    title = "Symulacja łańcucha Markowa w czasie ciągłym",
    x = "Czas",
    y = "Stan"
  ) +
  theme_minimal()
```



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 $s, t \in \mathbb{R}_+$ położmy 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 Feller'a 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

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

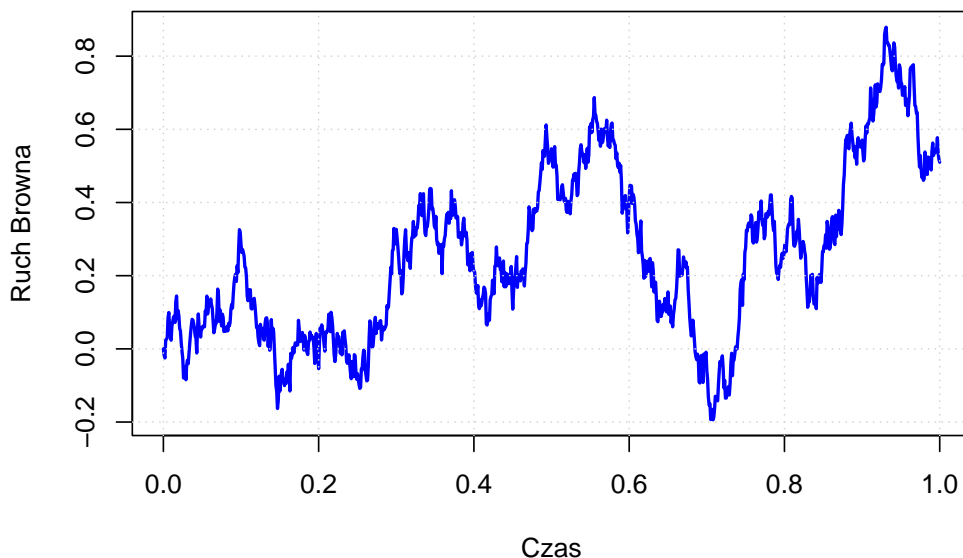
jest ciągle dla dowolnego n , dowolnych $t_1, \dots, t_n \in \mathbb{R}$ oraz dowolnych $f_1, \dots, f_n \in C_0(S)$.

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 ciągle.

Ruch Browna



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 $x \in S$ zdefiniujmy \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 $t \mapsto B_t + x$.

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 0.1.

Zadanie 0.6. Niech $S = \mathbb{Z}$. Pokaż, że łańcuch Markowa w czasie ciągłym (\mathbf{P}, \mathbb{F}) jest procesem Fellera 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 umotywować 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 Fellera 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 $C(S)$.

Działanie półgrupy Gaussa-Weierstrassa na funkcji $f(x)$, zdefiniowanej jako $\sin(x)$ na przedziale $[-\pi, \pi]$ i 0 poza nim. Funkcja wynikowa $T_t f(x)$ jest wyznaczana jako splot f z jądrem Gaussa dla zadanego parametru t .

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 $g_n = f_n - \lambda Lf_n$ 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:

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

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 $f - \lambda Lf = g = h - \lambda Lh$, 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 0.9 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 infinitesimalnym.

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 ciągle w t dla każdego x , a także ciągle 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 $(L, (L))$ jest generatorem infinitesimalnym. 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

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

Pochodna półgrupy Gaussa-Weierstrassa na funkcji $f(x)$, zdefiniowanej jako $\sin(x)$ na przedziale $[-\pi, \pi]$ i 0 poza nim. Funkcja wynikowa to $(T_t f(x) - f(x))/t$ dla zadanego parametru t .

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 0.9, 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 0.8. 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 ciągłe względem t , więc $T(t)f$ jest ciągłe 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 Fellera 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 trafenia $a > 0$ 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 Fellera. 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 infinitesimalnymi. 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 infinitesimalnym. 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 infinitesimalnego, 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 infinitezymalnym 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 Feller'a, 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 .

Wykład 6: od generatora przez półgrupę do procesu

2024-11-07

Piotr Dyszewski

W całym tym rozdziale $(L, \mathcal{D}(L))$ będzie generatorem infinitezymalnym. Naszym pierwszym zadaniem jest skonstruowanie odpowiadającej mu półgrupy Fellera. Aby to zrobić, wprowadzamy aproksymację L_ϵ zadaną przez

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

Zauważmy, że L_ϵ jest dobrze określone ponieważ

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

Ponadto, przy oznaczeniach jak wyżej,

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

, 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 0.13.

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

$$(I - \epsilon L)^{-1}f - \epsilon L_\epsilon f = f. \tag{19}$$

- Użyj części (a), aby pokazać, że L_ϵ jest generatorem infinytezymalny, oraz że $T_\epsilon(t)$ jest półgrupą Fellera, której generatorem jest L_ϵ .

Twierdzenie 0.13. 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 .

Proof. Najpierw sprawdzamy, że L_ϵ i L_δ komutują dla $\epsilon, \delta > 0$. Aby to sprawdzić zauważmy najpierw, że

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

co z kolei 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 . Korzystając teraz z (19) napiszmy

$$\begin{aligned} \epsilon\delta L_\epsilon L_\delta &= \left((I - \epsilon L)^{-1} - I\right) \left((I - \delta L)^{-1} - I\right) = \\ &= \left((I - \delta L)^{-1} - I\right) \left((I - \epsilon L)^{-1} - I\right) = \epsilon\delta L_\delta L_\epsilon. \end{aligned}$$

Zauważmy, że dzięki temu operatory L_δ , L_ϵ , $T_\epsilon(t)$ i $T_\delta(t)$ również komutują.

Chcąc pokazać, że $T_\epsilon(t)$ przy $\epsilon \rightarrow 0$ rzeczywiście zbiegają będziemy chcieli opierać się na zbieżności L_ϵ przy $\epsilon \rightarrow 0$. Uzasadnimy teraz nierówność, która uzasadni, że takie wnioskowanie jest możliwe. Stosując zasadnicze twierdzenie rachunku różniczkowego i całkowego otrzymujemy, że dla każdej $f \in C_0(S)$,

$$T_\epsilon(t)f - T_\delta(t)f = \int_0^t \frac{d}{ds} [T_\epsilon(s)T_\delta(t-s)]f \, ds = \int_0^t T_\epsilon(s)T_\delta(t-s)(L_\epsilon - L_\delta)f \, ds.$$

Skoro $\|T_\epsilon(s)\|, \|T_\delta(t-s)\| \leq 1$ udowodniona właśnie nierówność implikuje, że

$$\|T_\epsilon(t)f - T_\delta(t)f\| \leq t\|L_\epsilon f - L_\delta f\|.$$

Wystarczy pokazać zatem, że L_ϵ zbiegają. Spodziewamy się, że $(I - \epsilon L)^{-1} \rightarrow I$, co z kolei powinno implikować, że $L_\epsilon \rightarrow L$. Sprawdźmy teraz formalnie to rozumowanie. Dla $f \in \mathcal{D}(L)$ mamy

$$(I - \epsilon L)^{-1}f - f = \epsilon(I - \epsilon L)^{-1}Lf$$

Skoro $\|(I - \epsilon L)^{-1}\| \leq 1$, to

$$\|(I - \epsilon L)^{-1}f - f\| \leq \epsilon\|Lf\|.$$

W szczególności, dla $f \in \mathcal{D}(L)$,

$$\lim_{\epsilon \rightarrow 0} (I - \epsilon L)^{-1}f = f.$$

Skoro $(I - \epsilon L)^{-1}$ jest ciągle a $\mathcal{D}(L)$ gęste w $C_0(S)$, powyższa zbieżność ma miejsce dla wszystkich $f \in C_0(S)$. Ponieważ

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

dla $f \in \mathcal{D}(L)$, to

$$\lim_{\epsilon \rightarrow 0} L_\epsilon f = Lf \tag{20}$$

dla $f \in \mathcal{D}(L)$. Tak więc przez @ref, granica w (0.13) istnieje jednostajnie na ograniczonych zbiorach t dla $f \in \mathcal{D}$. Ponieważ zbiór $\mathcal{D}(L)$ jest gęsty w $C_0(S)$ i $T_\epsilon(t)$ jest kontrakcją, ten sam wniosek jest prawdziwy dla wszystkich $f \in C_0(S)$. Z udowodnionej właśnie zbieżności wynika, że a $T(t)$ spełnia własność półgrupy oraz wszystkie postulaty procesu Feller'a z wyjątkiem ostatniego.

Aby uzasadnić, że $(T_t)_{t \in \mathbb{R}_+}$ spełnia ostatnią własność pokażemy najpierw, że

$$(I - \alpha L)^{-1} = \alpha \int_0^\infty e^{-\alpha t} T_t dt.$$

Ustalmy w tym celu $\lambda > 0$ tak małe, że $\mathcal{R}(I - \lambda L) = C_0(S)$, i wybierzmy α tak, że $\alpha\lambda = 1$. Ustalmy $g \in C_0(S)$. Chcąc uzasadnić

$$(I - \alpha L)^{-1}g = \alpha \int_0^\infty e^{-\alpha t} T_t g dt$$

odwołamy się do aproksymacji L i $T(t)$ przez odpowiednio L_ϵ oraz $T_\epsilon(t)$. Niech

$$f_\epsilon = (I - \alpha L_\epsilon)^{-1}g = \alpha U_\epsilon(\alpha)g = \alpha \int_0^\infty e^{-\alpha t} T_\epsilon(t)g dt.$$

Z (0.13),

$$\lim_{\epsilon \rightarrow 0^+} f_\epsilon = f, \quad f = \alpha \int_0^\infty e^{-\alpha t} T(t)g dt.$$

Położmy $h_\epsilon = (I - \epsilon L)^{-1} f_\epsilon$. Wówczas z (19), $\lim_{\epsilon \rightarrow 0} h_\epsilon = f$. Niech wreszcie $h = (I - \lambda L)^{-1} g$. Naszym aktualnym celem () we wprowadzonej właśnie notacji jest pokazać, że $f = h$. Z definicji L_ϵ ,

$$\begin{aligned} \lambda L h_\epsilon &= L_\epsilon f_\epsilon = \alpha \int_0^\infty e^{-\alpha t} L_\epsilon T_\epsilon(t) f dt = \alpha \int_0^\infty e^{-\alpha t} \frac{d}{dt} T_\epsilon(t) f dt \\ &= -\alpha g + \alpha^2 \int_0^\infty e^{-\alpha t} T_\epsilon(t) f dt = \alpha(f_\epsilon - g) \rightarrow \alpha(f - g) \end{aligned}$$

gdy $\epsilon \rightarrow 0$. Dlatego,

$$\lim_{\epsilon \rightarrow 0} [(h_\epsilon - h) - \lambda L(h_\epsilon - h)] = 0,$$

z czego z kolei wynika, że $h_\epsilon - h \rightarrow 0$. Stąd $f = h$. Zastosowanie () do g_n i $f_n = (I - \lambda L)^{-1} g_n$ z Definicji 0.9 widzimy, że

$$\lim_{n \rightarrow \infty} \alpha \int_0^\infty e^{-\alpha t} (T(t) g_n)(x) dt = \lim_{n \rightarrow \infty} f_n(x) = 1$$

dla każdego $x \in S$. Niech $\mu_{t,x}$ będzie miarą zdefiniowaną przez

$$T(t)f(x) = \int f(y) \mu_{t,x}(dy).$$

Skoro $T(t)$ jest ograniczonym operatorem, to miara $\mu_{t,x}$ jest skończona. Jako, że $g_n \rightarrow 1$ punktowo, to twierdzenia o zbieżności ograniczonej mamy więc

$$T(t)g_n(x) \rightarrow \mu_{t,x}(S).$$

Stąd

$$1 = \alpha \int_0^\infty e^{-\alpha t} \mu_{t,x}(S) dt.$$

Innymi słowy transformata Laplace'a funkcji $t \mapsto \mu_{t,x}(S)$ jest $\alpha \rightarrow 1/\alpha$. Widzimy więc, że $\mu_{t,x}(S) \equiv 1$. Czyli $T(t)g_n \rightarrow 1$ punktowo. Oznacza to, że spełniona jest ostatnia własność w Definicji 0.9.

Wreszcie, sprawdzamy, że ta półgrupa Fellera $(T_t)_{t \in \mathbb{R}_+}$ ma generator L . Mamy

$$T_\epsilon(t)f - f = \int_0^t \frac{d}{ds} T_\epsilon(s) ds = \int_0^t T_\epsilon(s) L_\epsilon f ds.$$

Jeśli $f \in \mathcal{D}(L)$, to (0.13), (20) oraz własność kontrakcji $T_\epsilon(t)$ implikują, że można przejść do granicy w tym równaniu, aby uzyskać

$$T(t)f - f = \int_0^t T(s) L f ds.$$

Zatem

$$\lim_{t \rightarrow 0} \frac{T(t)f - f}{t} = Lf \text{ dla } f \in \mathcal{D}(L).$$

Pokazaliśmy właśnie, że

$$\mathcal{D}(L) \subseteq \mathcal{D}'(L) = \left\{ f \in C_0(S) : \lim_{t \rightarrow 0} (T(t)f - f)/t \text{ istnieje} \right\}.$$

Wiemy, że obie pary $(L, \mathcal{D}(L))$ oraz $(L, \mathcal{D}'(L))$ są generatorami infinitezymalnymi. Niech $f \in \mathcal{D}'(L)$. Rozważmy $g = (I - \lambda L)f$. Istnieje $h \in \mathcal{D}(L)$ takie, że $g = (I - \lambda L)h$. Stąd $(I - \lambda L)(f - h) = 0$ a co za tym idzie $f = h \in \mathcal{D}(L)$. Pokazaliśmy właśnie, że $\mathcal{D}(L) = \mathcal{D}'(L)$. \square

Twierdzenie 0.14. *Jeśli $(T_t)_{t \geq 0}$ jest półgrupą Fellera, to istnieje proces Fellera (\mathbf{P}, \mathbb{F}) spełniający*

$$\mathbf{E}_x f(X(t)) = T_t f(x)$$

dla $x \in S, t \geq 0$, oraz $f \in C_0(S)$.

Proof. Ustalmy $x \in S$. Pokażemy najpierw, że istnieje proces $Y(t), t \in \mathbb{Q}^+$ na pewnej przestrzeni probabilistycznej, który ma pożądane rozkłady skończenie wymiarowe oraz $Y(0) = x$.

Dla dowolnego $n \in \mathbb{N}$ i dowolnych $0 \leq t_1 < t_2 < \dots < t_n$ definiujemy rozkład μ_{t_1, \dots, t_n} na S^n poprzez następujący wymóg: dla dowolnych f_1, \dots, f_n z $C_0(S)$ zachodzi

$$\int f_1(y_1) f_2(y_2) \cdots f_n(y_n) \mu_{t_1, \dots, t_n}(dy_1, \dots, dy_n) = T_{t_1}(f_1 T_{t_2 - t_1}(f_2 T_{t_3 - t_2}(\dots)))(x). \quad (21)$$

Funkcje o rozdzielonych zmiennych $f_1(y_1) f_2(y_2) \cdots f_n(y_n)$ są liniowo gęste w $C(S^n)$. Miara μ_{t_1, \dots, t_n} jest więc dobrze określona (istnieje dokładnie jedna spełniająca zadany warunek). Aby powołać się na twierdzenie Kołmogorowa o istnieniu procesu należy uzasadnić, że skonstruowane w ten sposób miary są zgodne. Dokładniej, że dla dowolnego $j \leq n$, i dowolnych borelowskich $A_1, \dots, A_n \in S$,

$$\begin{aligned} \mu_{t_1, \dots, t_n}(A_1 \times \cdots \times A_{j-1} \times S \times A_{j+1} \times \cdots \times A_n) = \\ \mu_{t_1, \dots, t_{j-1}, t_{j+1}, \dots, t_n}(A_1 \times \cdots \times A_{j-1} \times A_{j+1} \times \cdots \times A_n). \end{aligned}$$

Wynika to z (21) dla $f_j \equiv 1$ i zastosowaniu własności półgrupy $(T_t)_{t \in \mathbb{R}_+}$. Zastosowanie twierdzenia Kołmogorowa pozwala wywnioskować istnienie procesu o zadanych rozkładach skończenie wymiarowych. Jednakże nie daje ono żadnej kontroli nad regularnością trajektorii. Dlatego jesteśmy zmuszeni przeprowadzić dodatkową konstrukcję.

Stosując konstrukcję w poprzedniego paragrafu dla wymiernych t otrzymujemy proces $(Y_q^{(x)} : q \in \mathbb{Q}_+)$ o rozkładach zadanych rozkładach skończenie wymiarowych:

$$\mathbb{P}[Y_{t_1}^{(x)} \in A_1, \dots, Y_{t_n}^{(x)} \in A_n] = \mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n)$$

oraz taki, że $\mathbb{P}[Y_0^{(x)} = x] = 1$. Proces ten ma własność Markowa. Z definicji miary μ_{t_1, \dots, t_n} ,

$$\mathbb{E}[f_1(Y_{t_1}^{(x)}) \cdots f_n(Y_{t_n}^{(x)})] = \mathbb{E}[f_1(Y_{t_1}^{(x)}) \cdots f_{n-1}(Y_{t_{n-1}}^{(x)}) T_{t_n - t_{n-1}} f_n(Y_{t_n}^{(x)})]$$

Powyższe implikuje

$$\mathbb{E}[f_n(Y_{t_n}^{(x)}) | Y_q^{(x)}, q \leq t_{n-1}] = T_{t_n - t_{n-1}} f_n(Y_{t_{n-1}}^{(x)})$$

W analogiczny sposób uzasadniamy własność Markowa.

Jeśli $f \in C_0(S)$ jest nieujemna, to

$$e^{-\alpha t} T(t) U(\alpha) f = \int_t^\infty e^{-\alpha s} T(s) f ds \leq U(\alpha) f.$$

Stąd wynika, że $e^{-\alpha t} U(\alpha) f(Y(t))$ jest (ograniczonym) supermartingąłem względem filtracji $\{\mathcal{F}_t, t \in \mathbb{Q}^+\}$, gdzie \mathcal{F}_t jest generowana przez

$$(Y(s), s \in \mathbb{Q} \cap [0, t]).$$

Rzeczywiście mamy

$$\mathbb{E}[e^{-\alpha(t+s)} U(\alpha) f(Y_{t+s}) | \mathcal{F}_s] = e^{-\alpha(t+s)} T_t U(\alpha) f(Y_s) \leq e^{-\alpha s} U(\alpha) f(Y_s).$$

Wówczas prawdopodobieństwem jeden, prawostronne i lewostronne granice tego supermartingąłu wzdłuż liczb wymiernych istnieją w każdym $t \in [0, \infty)$.

Dowód twierdzenia o zbieżności nadmartingąłów sprowadza się pokazania, że w prawdopodobieństwie jeden liczba przejść w dół przez każdy przedział

o końcach wymiernych jest skończona. Każda funkcja zmiennej rzeczywistej o tej własności musi mieć lewostronną i prawostronną granicę w każdym punkcie.

Dla każdej α i każdej nieujemnej f proces $e^{-\alpha t}U(\alpha)f(Y_t)$ ma lewostronne i prawostronne granice. Oczywiście, zbiór o prawdopodobieństwie zero, na którym to się nie udaje, może zależeć od f i α . Musimy więc ograniczyć naszą uwagę do gęstego, przeliczalnego zbioru f i α , aby stwierdzić istnienie prawostronnych i lewostronnych granic procesu $Y(s)$ samego w sobie. Ponieważ $\alpha U(\alpha)f \rightarrow f$ gdy $\alpha \rightarrow \infty$, oraz $C_0(S)$ jest ośrodkową przestrzenią metryczną, możemy po prostu użyć funkcji $U(\alpha)f$, gdzie α jest liczbą całkowitą, oraz f jest wzięte z pewnego przeliczalnego gęstego zbioru w $C_0(S)$.

Istnieje nadal jeden problem, że te prawostronne i lewostronne granice mogą być jednopunktową kompaktyfikacją S , i wykluczenie tego przypadku wymaga dodatkowego argumentu. Chodzi o to, żeby uzasadnić, że Y nie ucieka do nieskończoności. Niech $f \in C_0(S)$ będzie ściśle dodatnie. Rozważmy supermartingał $M(t) = e^{-tU(1)f(Y(t))}$. To jest ściśle dodatnie, więc $\inf_{s \in \mathbb{Q} \cap [0, t]} M(s) > 0$ prawie na pewno dla każdego $t > 0$ (zadanie). Oznacza to, że z prawdopodobieństwem jeden Y_s dla $s \leq t$ są zawarte w zbiorze zwartym.

Wiemy teraz, że z prawdopodobieństwem jeden, $Y(s)$ ma prawostronne i lewostronne granice w każdym $t \in [0, \infty)$. Więc możemy zdefiniować

$$X(t) = \lim_{\substack{s \downarrow t \\ s \in \mathbb{Q}}} Y(s).$$

To jest automatycznie prawostronnie ciągłe i ma lewostronne granice. Ponieważ skończenie wymiarowe rozkłady odpowiadające $T(t)$ są słabo ciągłe, $X(t)$ ma poprawne skończenie wymiarowe rozkłady. \square

Wykład 7: Rozkłady stacjonarne, zaburzenia ruchu Browna

2024-11-14

Piotr Dyszewski

Rozkłady stacjonarne

Interesować nas będzie asymptotyczne zachowanie procesów Feller'a. Podobnie jak w przypadku łańcuchów Markowa w czasie dyskretnym rozkłady graniczne są niezmiennicze ze względu na funkcje przejścia. Przez Σ oznaczać będziemy σ -ciało zbiorów borelowskich S , czyli najmniejsze σ -ciało zawierające wszystkie otwarte podzbiory S . Skoro S jest ośrodkowa, to Σ jest generowane przez wszystkie kule otwarte.

Dla Procesu Feller'a (\mathbf{P}, \mathbb{F}) oraz rozkładu prawdopodobieństwa μ na S definiujemy miarę probabilistyczną \mathbf{P}_μ na (S, Σ) wzorem

$$\mathbf{P}_\mu[A] = \int_S \mathbf{P}_x[A] \mu(dx), \quad A \in \mathcal{F}.$$

W tym miejscu zachęcamy czytelnika do sprawdzenia, że odwzorowanie $x \mapsto \mathbf{P}_x[A]$ jest mierzalne dla $A \in \mathcal{F}$. Miara \mathbf{P}_μ to rozkład procesu Markowa przy rozkładzie początkowym μ .

Definicja 0.10. Niech (\mathbf{P}, \mathbb{F}) będzie procesem Feller'a. Rozkład prawdopodobieństwa π na (S, Σ) nazywamy rozkładem stacjonarnym jeżeli

$$\mathbf{P}_\pi[X(t) \in A] = \pi(A)$$

dla każdego $A \in \Sigma$.

Chcielibyśmy wiedzieć, jak określić na podstawie generatora, czy miara prawdopodobieństwa na S jest stacjonarna dla procesu Fellera. Z tego powodu przepiszemy powyższą definicję w terminach półgrupy. Jeśli μ jest miarą prawdopodobieństwa na S , rozkład procesu w czasie t , gdy rozkład początkowy jest μ , oznaczamy przez $\mu T(t)$. Spełnia on zależność

$$\int f d(\mu T(t)) = \int T(t)f d\mu = \mathbf{E}_\mu[f(X(t))]$$

dla $f \in C_0(S)$. Tutaj \mathbf{E}_μ to wartość oczekiwana odpowiadająca \mathbf{P}_μ . Równoważnie

$$\mathbf{E}_\mu[Y] = \int_X \mathbf{E}_x[Y] \mu(dx)$$

dla każdej ograniczonej zmiennej losowej $Y: \Omega \rightarrow \mathbb{R}$.

Definicja 0.11. Miara prawdopodobieństwa μ na S jest stacjonarna dla procesu Fellera z półgrupą $T(t)$, jeśli $\mu T(t) = \mu$ dla wszystkich $t \geq 0$, tzn. jeśli

$$\int T(t)f d\mu = \int f d\mu \quad \text{dla wszystkich } f \in C_0(S) \text{ i } t \geq 0.$$

Zadanie 0.14. Pokaż, że jeśli μ jest miarą prawdopodobieństwa na S i $\mu T(t) \Rightarrow \nu$, to ν jest stacjonarna.

Twierdzenie 0.15. *Miara prawdopodobieństwa μ na S jest stacjonarna dla odpowiadającego procesu wtedy i tylko wtedy, gdy*

$$\int Lf d\mu = 0 \quad \text{dla wszystkich } f \in D.$$

Proof. Przypuśćmy, że μ jest stacjonarna, i weźmy $f \in \mathcal{D}(L)$. Wtedy

$$\int Lf d\mu = \int \lim_{t \rightarrow 0} \frac{T(t)f - f}{t} d\mu = \lim_{t \rightarrow 0} \frac{\int T(t)f d\mu - \int f d\mu}{t} = 0.$$

Przeciwnie, przypuśćmy, że $\int Lf d\mu = 0$ dla wszystkich $f \in \mathcal{D}(L)$ i jeśli $f \in \mathcal{D}(L)$ oraz $f - \lambda Lf = g$, to $\int f d\mu = \int g d\mu$. Iterując to, otrzymujemy

$$\int (I - \lambda L)^{-n} g d\mu = \int g d\mu.$$

Biorąc $\lambda = t/n$ i przechodząc z $n \rightarrow \infty$ wnioskujemy, że

$$\int T(t)g \, d\mu = \int g \, d\mu.$$

□

Oto wystarczający warunek na istnienie rozkładu stacjonarnego.

Twierdzenie 0.16. *Jeśli S jest przestrzenią zwartą, to istnieje miara stacjonarna.*

Proof. Rozważmy proces Feller'a z dowolnym rozkładem początkowym μ . Niech ν_n będzie rozkładem zmiennej Z definiujemy miarę ν_n na S poprzez warunek

$$\int_S f(y) \nu_n(dy) = \mathbb{E} [\mathbf{E}_\mu [f(X_{nU})]] = \mathbb{E} \left[\int_S T_{nU} f(y) \mu(dy) \right].$$

Dla $f \in C_0(S)$, własność półgrupy daje

$$\int T(t)f(y) \nu_n(dy) = \mathbb{E} \left[\int_S T_{nU+t} f(y) \mu(dy) \right].$$

także

$$\begin{aligned} \int f \, d\nu_n - \int T(t)f \, d\nu_n &= \int f \, d\nu_n - \int f \, d(\nu_n T(t)) \\ &= \frac{1}{n} \left[\int_0^t \int_S T(r)f \, d\mu \, dr - \int_n^{n+t} \int_S T(r)f \, d\mu \, dr \right]. \end{aligned} \quad (22)$$

Prawa strona (22) dąży do zera gdy $n \rightarrow \infty$.

Teraz, ponieważ S jest zwarty, twierdzenie Prochorowa, implikuje, że istnieje podciąg ν_{n_k} taki, że

$$\nu_{n_k} \Rightarrow \nu$$

dla pewnej miary prawdopodobieństwa ν na S . Zatem, ponieważ $T(t)f \in C(S)$, możemy przejść do granicy w (22) wzdłuż ciągu ν_{n_k} , aby otrzymać

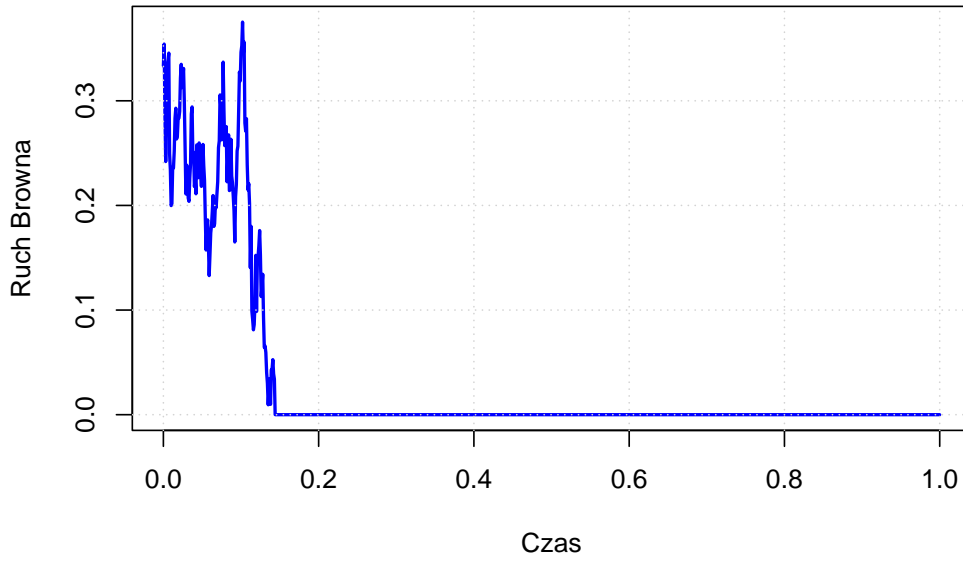
$$\int f \, d\nu = \int T(t)f \, d\nu.$$

Ponieważ to zachodzi dla wszystkich $f \in C_0(S)$, wynika stąd, że $\nu T(t) = \nu$. □

Zaburzenia ruchu Browna

Przykład 0.5. Rozważmy ruch Browna na $[0, \infty)$ z absorpcją w 0.

Ruch Browna zapoczątkowany w 1/3 z absorpcją w 0



Niech τ będzie czasem pierwszego uderzenia w 0. Zdefiniujmy

$$X_a(t) = \begin{cases} X(t) & \text{jeśli } t < \tau, \\ 0 & \text{jeśli } t \geq \tau, \end{cases}$$

oraz oznaczmy przez L_a i $T_a(t)$ odpowiednio generator i półgrupę. Dla $f \in C_0[0, \infty)$, niech f_o będzie „nieparzystym” przedłużeniem f na \mathbb{R} :

$$f_o(x) = \begin{cases} f(x) & \text{jeśli } x \geq 0, \\ 2f(0) - f(-x) & \text{jeśli } x < 0. \end{cases}$$

Z zasady odbicia dla każdej $g \in C_0[0, \infty)$,

$$\mathbf{E}_x [g(X(t)) \mathbf{1}_{\{t \geq \tau\}}] = \mathbf{E}_x [g(-X(t)) \mathbf{1}_{\{t \geq \tau\}}].$$

Biorąc $g = f_o$,

$$\mathbf{E}_x [f_o(X(t)) \mathbf{1}_{\{t \geq \tau\}}] = \mathbf{E}_x [f_o(-X(t)) \mathbf{1}_{\{t \geq \tau\}}].$$

Obie te wielkości są równe

$$\frac{1}{2} \mathbf{E}_x \left[(f_o(X(t)) + f_o(-X(t))) \mathbf{1}_{\{t \geq \tau\}} \right].$$

Ostatnie wyrażenie, z definicji f_o jest równe

$$f(0) \mathbf{P}_x(t \geq \tau).$$

Podsumowując dla $x \geq 0$,

$$T_a(t)f(x) = \mathbf{E}_x \left[f(X(t)) \mathbf{1}_{\{t < \tau\}} \right] + f(0) \mathbf{P}_x(t \geq \tau) = \mathbf{E}_x f_o(X(t)).$$

Oczywiście $f_o \notin C(\mathbb{R})$ o ile $f(0) = 0$. Niemniej jednak, skoro

$$f_o''(x) = \begin{cases} f''(x) & \text{jeśli } x > 0, \\ -f''(-x) & \text{jeśli } x < 0, \end{cases}$$

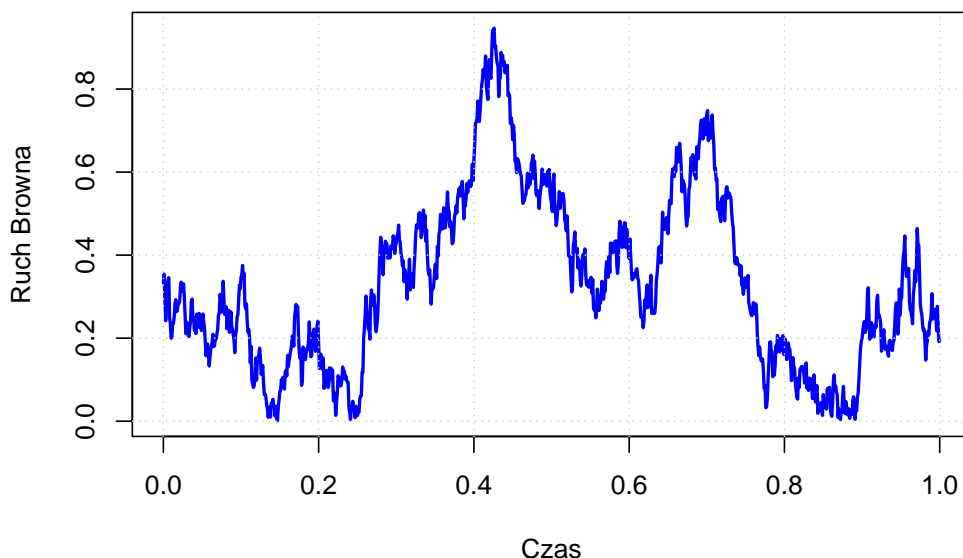
wtedy $f''(0) = 0$ jest potrzebne, aby f_o'' było ciągłe. Wynika z tego, że

$$\mathcal{D}(L_a) = \{f \in C_0[0, \infty) : f'' \in C[0, \infty), f''(0) = 0\},$$

a dla $f \in \mathcal{D}(L_a)$, $L_a f = \frac{1}{2} f''$.

Przykład 0.6. Rozważmy ruch Browna na $[0, \infty)$ z odbiciem w 0.

Ruch Browna zapoczątkowany w 1/3 z odbiciem w 0



Proces ten jest zdefiniowany jako

$$X_r(t) = |X(t)|,$$

a jego generator i półgrupa będą oznaczane odpowiednio przez L_r i $T_r(t)$.
Jeśli $f \in C_0[0, \infty)$, niech f_e będzie parzystym przedłużeniem f na \mathbb{R} :

$$f_e(x) = \begin{cases} f(x) & \text{jeśli } x \geq 0, \\ f(-x) & \text{jeśli } x < 0. \end{cases}$$

Wtedy

$$T_r(t)f(x) = \mathbf{E}_x[f(|X(t)|)] = \mathbf{E}_x f_e(X(t)) \quad \text{dla } x \geq 0.$$

Zatem,

$$f \in \mathcal{D}(L_r) \iff f_e \in \mathcal{D}(L).$$

Wynika z tego, że

$$\mathcal{D}(L_r) = \{f \in C[0, \infty) : f', f'' \in C[0, \infty), f'(0) = 0\},$$

a dla $f \in \mathcal{D}(L_r)$, $L_r f = \frac{1}{2} f''$.

Przykład 0.7. Zaprezentujemy teraz ruch Browna na $[0, \infty)$ z lepkiem 0. Dla $c > 0$, rozważmy operator L_c zdefiniowany jako $L_c f = \frac{1}{2} f''$ na

$$\mathcal{D}(L_c) = \{f \in C_0[0, \infty) : f'' \in C[0, \infty), f'(0) = c f''(0)\}.$$

Zauważmy, że graniczne przypadki $c \downarrow 0$ i $c \uparrow \infty$ odpowiadają odpowiednio odbiciu i absorpcji w 0. Jest to generator prawdopodobieństwa — dowód jest pozostawiony jako ćwiczenie. Oto weryfikacja własności (d) w Definicji 2: Dla $g \in C_0[0, \infty)$ i $\lambda > 0$, musimy rozwiązać $f - \lambda L_c f = g$ dla $f \in \mathcal{D}(L_c)$. Niech $f_a \in \mathcal{D}(L_a)$ oraz $f_r \in \mathcal{D}(L_r)$ będą rozwiązaniami

$$f_a - \lambda L_a f_a = g \quad \text{oraz} \quad f_r - \lambda L_r f_r = g.$$

Ponieważ wszystkie trzy generatory są równe $\frac{1}{2} f''$ na swoich dziedzinach,

$$f = \gamma f_a + (1 - \gamma) f_r$$

jest wymaganym rozwiązaniem, pod warunkiem że $f'(0) = c f''(0)$. Ma to miejsce, gdy γ spełnia

$$\gamma f'_a(0) = c(1 - \gamma) f''_r(0).$$

Aby znaleźć wartość γ , $f'_a(0)$ i $f''_r(0)$ muszą mieć ten sam znak. Aby to sprawdzić, rozważmy $h = f_a - f_r$. Wtedy $h - \frac{\lambda}{2}h'' \equiv 0$, więc, ponieważ h jest ograniczone,

$$h(x) = h(0)e^{-x\sqrt{2/\lambda}}.$$

Wynika z tego, że

$$f'_a(0) = -\sqrt{2/\lambda}h(0)$$

oraz

$$f''_r(0) = -\left(\frac{2}{\lambda}\right)h(0),$$

więc mają ten sam znak, i

$$\gamma = \frac{2c}{2c + \sqrt{2\lambda}}.$$

Aby powiedzieć coś o zachowaniu tego procesu, gdy odwiedza 0, napiszmy

$$f = \alpha U_c(\alpha)g,$$

gdzie $\alpha = \lambda^{-1}$, a U_c jest rozwiązaniem dla procesu $X_c(t)$ z generatorem L_c . Można to zapisać jako

$$f(x) = \frac{2cf_a(x) + \sqrt{2\lambda}f_r(x)}{2c + \sqrt{2\lambda}} = \alpha \int_0^\infty e^{-\alpha t} \mathbf{E}_x g(X_c(t)) dt.$$

Zastosujmy tę tożsamość do ciągu funkcji g , które są nieujemne i rosną do 1_(0,∞). Odpowiadające im f , f_a , oraz f_r rosną odpowiednio do

$$\alpha \int_0^\infty e^{-\alpha t} \mathbf{P}_x(X_c(t) > 0) dt, \quad \alpha \int_0^\infty e^{-\alpha t} \mathbf{P}_x(X_a(t) > 0) dt,$$

oraz 1. Biorąc $x = 0$, otrzymujemy

$$\mathbf{E}_0 \int_0^\infty \alpha e^{-\alpha t} 1_{\{X_c(t) > 0\}} dt = \frac{1}{1 + c\sqrt{2\alpha}}.$$

Zatem miara Lebesgue'a zbioru $\{t \geq 0 : X_c(t) = 0\}$ jest dodatnia, w przeciwieństwie do przypadku procesu odbijanego, który odpowiada $c = 0$.

Wykład 8: Układy spinowe

2024-11-21

Piotr Dyszewski

Motywacje

Aby umotywować nasze przyszłe działania przedyskutujemy kilka przykładów. Od tej pory niech $G = (V, E)$ będzie przeliczalnym grafem prostym o ograniczonym stopniu. Dokładniej zakładamy, że zbiór jego wierzchołków V jest skończony bądź przeliczalny oraz, że

$$\sup_{x \in V} \deg(x) < \infty.$$

Rozważmy następujące trzy procesy na G .

Przykład 0.8. (Voter model). Przypuśćmy, że na G są dwie wzajemnie zwalczające się frakcje (można myśleć o republikanach i demokratkach). Na każdy wierzchołek jednej frakcji oddziałują (poprzez indoktrynację) sąsiedzi z frakcji przeciwnej. Na skutek czego niektóre wierzchołki zmieniają frakcję na przeciwną. Dokładniej każdy $x \in V$ zmienia frakcję w intensywnością równą

$$c(x) = \#\{y \in V \mid x \sim y, y \text{ jest z innej frakcji niż } x\}.$$

Jest to równoważne z następującym opisem.

- Dla każdej pary sąsiednich wierzchołków $x \sim y$ pochodzących z różnych frakcji losujemy niezależnie dwie liczby $E_{x \rightarrow y}$ oraz $E_{y \rightarrow x}$ z rozkładu wykładniczego z parametrem jeden. $E_{x \rightarrow y}$ interpretujemy jako czas, jaki potrzebuje x aby przekabacić y .

- W momencie, w którym którykolwiek z wierzchołków przeciągnie sąsiada (powiedzmy y) na swoją stronę, losujemy nowe wagi dla sąsiadów y (według nowego układu obu frakcji).

Oczywiście powyższa, naiwna konstrukcja ma sens tylko, gdy graf G jest skończony. W przypadku grafów nieskończonych wymagana jest odpowiednia konstrukcja, którą opiszemy niebawem.

Przykład 0.9. (Contact process). Załóżmy, że na grafie G panuje epidemia. Niech $\lambda > 0$ będzie ustalonym parametrem. Proces rozwija się według następujących zasad.

- Chore wierzchołki niezależnie zarażają swoich zdrowych sąsiadów z czasem wykładniczym z parametrem λ .
- Chore wierzchołki zdrowieją niezależnie z czasem wykładniczym z parametrem jeden.

Ponownie powyższy opis ma sens jedynie w przypadku, gdy G jest skończony. Dla nieskończonych grafów G będziemy posługiwać się równoważną charakterystyką

- Każdy chory wierzchołek zdrowieje z intensywnością 1
- Każdy zdrowy wierzchołek x choruje z intensywnością

$$\lambda \# \{y \sim x : y \text{ chory}\}.$$

Przykład 0.10. (Exclusion process). Rozważmy zjawisko migracji na grafie G . Załóżmy, że pewne wierzchołki G są zajmowane przez osobników ustalonego gatunku (przykładowo krowy). Każdy osobnik czeka niezależnie wykładniczy czas z parametrem jeden, po czym podejmuje próbę przemieszczenia się.

- Jeżeli osobnik zajmuje wierzchołek $x \in V$, to losuje wierzchołek y z prawdopodobieństwem $p(x, y)$.
- Jeżeli wierzchołek y nie jest zajęty, to osobnik przechodzi do y .

W każdym z powyższych przykładów stan układu w chwili t można zakodować przy pomocy $\eta_t \in \{0, 1\}^V$. Okazuje się, że wówczas $\eta = (\eta_t)_{t \in \mathbb{R}_+}$ staje się procesem Feller'a. Aby się o tym dokładnie przekonać musimy przeanalizować generatory wynikające z przykładów.

Systemy spinowe

Właściwością, która odróżnia systemy spinowe od innych procesów Fellera na $\{0, 1\}^S$, jest to, że poszczególne przejścia obejmują tylko jedną lokalizację. Niech $c: V \times \{0, 1\}^V \rightarrow \mathbb{R}_+$ będzie ograniczoną funkcją taką, że dla każdego $x \in V$, $c(x, \cdot): \{0, 1\}^V \rightarrow \mathbb{R}_+$ jest ciągła. $c(x, \eta)$ interpretować będziemy jako intensywność, z jaką stan η zmienia się poprzez zmianę wartości w x . Dla $x \in V$ oraz $\eta \in \{0, 1\}^V$ definiujemy $\eta^{(x)} \in \{0, 1\}^V$ wzorem

$$\eta^{(x)}(y) = \begin{cases} \eta(y), & y \neq x \\ 1 - \eta(x), & y = x \end{cases}.$$

Dla f pochodzącego z odpowiedniego podzbioru $C_0(\{0, 1\}^V)$ chcemy położyć

$$Lf(\eta) = \sum_{x \in V} c(x, \eta) [f(\eta^{(x)}) - f(\eta)]. \quad (23)$$

Okazuje się, że dokładne napisanie dziedziny jest problematyczne. Aby obejść tę trudność rozważmy

$$D = \left\{ f \in C(\{0, 1\}^V) : \|f\|_o := \sup_{\eta} \sum_x |f(\eta^{(x)}) - f(\eta)| < \infty \right\}. \quad (24)$$

Naszym celem jest pokazanie, że zdefiniowanie L na D wystarcza do zdefiniowania procesu.

Dygresja analityczna

Generator infinitesimalny nie należy do najprostszych obiektów w teorii procesów Fellera. Jednym z powodów jest konieczność uwzględnienia dziedziny która, jak się już przekonaaliśmy, ma istotny wpływ na kształt generowanego procesu. Rzadko się jednak zdarza, że dziedzinę można opisać jawnie. Dlatego często definiuje się proponowany generator na wygodnej podprzestrzeni dziedziny, a następnie bierze się domknięcie.

Definicja 0.12. Operator liniowy $(L, \mathcal{D}(L))$ na $C_0(S)$ nazywany jest domkniętym, jeśli jego wykres

$$\Gamma(L) = \{(f, Lf) : f \in \mathcal{D}(L)\}$$

jest domkniętym podzbiorem $C_0(S) \times C_0(S)$.

Operator L jest domknięty wtedy i tylko wtedy, gdy $f_n \in \mathcal{D}(L)$ są takie, że $f_n \rightarrow f$ oraz $Lf_n \rightarrow h$, to $f \in \mathcal{D}(L)$ oraz $h = Lf$.

Definicja 0.13. Operator liniowy $(L, \mathcal{D}(L))$ nazywany jest domykającym, jeżeli domknięcie jego wykresu $\Gamma(L)$ jest wykresem operatora liniowego. W takiej sytuacji definiujemy domknięcie \bar{L} operatora L poprzez

$$\Gamma(\bar{L}) = \overline{\Gamma(L)}.$$

Operator L jest domykający wtedy i tylko wtedy, gdy $f_n \rightarrow 0$ oraz $Lf_n \rightarrow h$ implikują $h = 0$. Nie każdy operator liniowy ma domknięcie. Na przykład, przypuśćmy, że $S = [0, 1]$ i

$$\mathcal{D}(L) = \{f \in C(S) : f'(0) \text{ istnieje}\} \quad \text{ i } \quad Lf(x) = f'(0)x \text{ dla } f \in \mathcal{D}(L).$$

Wtedy domknięcie wykresu L nie jest wykresem operatora liniowego. Jednakże w kontekście Definicji 0.9 taka sytuacja nie występuje.

Fakt 0.1. Niech $(L, \mathcal{D}(L))$ będzie operatorem liniowym na $C_0(S)$.

- Przypuśćmy, że L spełnia (GI1)-(GI3) z Definicji 0.9. Wtedy L jest domykający, a jego domknięcie spełnia (GI1)-(GI3).
- Jeśli L spełnia (GI1)-(GI4) z Definicji 0.9, wtedy L jest domknięty.
- Jeśli L spełnia (GI3) i (GI4), to

$$\mathcal{R}(I - \lambda L) = C_0(S) \quad \text{ dla każdego } \lambda > 0.$$

- Jeśli L jest domknięty i spełnia (GI3), to $\mathcal{R}(I - \lambda L)$ jest domkniętym podzbiorem $C_0(S)$.

Proof. Dla pierwszego stwierdzenia, musimy udowodnić, że $f_n \in \mathcal{D}(L)$, $f_n \rightarrow 0$, oraz $Lf_n \rightarrow h$ implikuje $h = 0$. Aby to zrobić, wybierzmy $g \in \mathcal{D}(L)$. Korzystając z (12),

$$\|(I - \lambda L)(f_n + \lambda g)\| \geq \|f_n + \lambda g\|, \quad \lambda > 0.$$

Biorąc $n \rightarrow \infty$ i następnie dzieląc przez λ , dostajemy

$$\|g - h - \lambda Lg\| \geq \|g\|.$$

Jeżeli teraz wybierzemy $\lambda \rightarrow 0$, to otrzymamy

$$\|g - h\| \geq \|g\|.$$

Skoro $g \in \mathcal{D}(L)$ jest wzięte z gęstego zbioru otrzymujemy $h = 0$. Domknięcie \bar{L} spełnia własności (GI1) i (GI2), ponieważ jest rozszerzeniem L . Aby sprawdzić, czy spełnia własność (GI3), przypuśćmy, że $f \in \mathcal{D}(\bar{L})$, $\lambda \geq 0$ i $f - \lambda \bar{L}f = g$. Przez definicję domknięcia, istnieją $f_n \in \mathcal{D}(L)$, takie że $f_n \rightarrow f$ i $Lf_n \rightarrow \bar{L}f$. Przez własność (GI3) dla L ,

$$\inf_{x \in S} f_n(x) \geq \inf_{x \in S} g_n(x),$$

gdzie $g_n = f_n - \lambda Lf_n$. Teraz niech $n \rightarrow \infty$. Dostajemy

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

czyli własność (GI3) dla \bar{L} .

Dla dowodu drugiej części faktu, niech \bar{L} będzie domknięciem L . Jeśli $f \in \mathcal{D}(\bar{L})$ i $\lambda > 0$ jest małe, przez własność (GI4) istnieje $h \in \mathcal{D}(L)$ takie że

$$h - \lambda Lh = f - \lambda \bar{L}f, \quad (25)$$

czyli $(h - f) - \lambda \bar{L}(h - f) = 0$. Z (12), $h = f$, a wtedy $\bar{L}f = Lh$ z (25). Więc, $\bar{L} = L$ zgodnie z tezą.

Przechodząc do trzeciego stwierdzenia, wystarczy sprawdzić, że dla $0 < \lambda < \gamma$ warunek $\mathcal{R}(I - \lambda L) = C_0(S)$ implikuje $\mathcal{R}(I - \gamma L) = C_0(S)$. Załóżmy, że $g \in C_0(S)$, i zdefiniujmy $\Gamma : C(S) \rightarrow \mathcal{D}(L)$ przez

$$\gamma \Gamma h = \lambda(I - \lambda L)^{-1}g + (\gamma - \lambda)(I - \lambda L)^{-1}h.$$

Definicja ta jest poprawna, ponieważ założyliśmy $\mathcal{R}(I - \lambda L) = C_0(S)$. Mamy

$$\gamma \|\Gamma h_1 - \Gamma h_2\| = (\gamma - \lambda) \|(I - \mathcal{L})^{-1}(h_1 - h_2)\| \leq (\gamma - \lambda) \|h_1 - h_2\|.$$

Stąd Γ jest odwzorowaniem zwężającym, a więc z Twierdzenia Banacha o punkcie stałym posiada jedyny punkt stały f . Wówczas $f \in \mathcal{D}(L)$ oraz

$$\gamma(I - \lambda \mathcal{L})f = \lambda g + (\gamma - \lambda)f.$$

Co można przekształcić do postaci $f - \gamma Lf = g$. Czyli $g \in \mathcal{R}(I - \gamma L)$, co należało uzasadnić.

Aby udowodnić ostatnie stwierdzenie, załóżmy, że $g_n \in \mathcal{R}(I - \lambda L)$ i $g_n \rightarrow g$. Wtedy możemy zdefiniować $f_n \in \mathcal{D}(L)$ przez

$$f_n - \lambda Lf_n = g_n. \quad (26)$$

Wówczas

$$(f_n - f_m) - \lambda L(f_n - f_m) = g_n - g_m,$$

a zatem $\|f_n - f_m\| \leq \|g_n - g_m\|$. Ponieważ $\{g_n\}_{n \in \mathbb{N}}$ jest ciągiem Cauchy'ego, to $\{f_n\}_{n \in \mathbb{N}}$ również. Niech $f = \lim_{n \rightarrow \infty} f_n$. Ponieważ $f_n \rightarrow f$ i $g_n \rightarrow g$, to z (26) wynika, że $\lim_{n \rightarrow \infty} Lf_n$ również istnieje. Ponieważ L jest domknięte, granicą jest Lf , a więc $f - \lambda Lf = g$, co oznacza, że $g \in \mathcal{R}(I - \lambda L)$, czego należało dowieść. \square

Konstrukcja systemów spinowych

Naszym pierwszym celem jest znalezienie naturalnych warunków na $c(x, \eta)$, które gwarantują, że domknięcie (L, D) , gdzie L i D są dane odpowiednio przez (23) oraz (24) jest operatorem infinitezymalnym. Warunki (GI1), (GI2), (GI3) i (GI5) są łatwe do sprawdzenia i nie wymagają dodatkowych założeń. Prawdziwym wyzwaniem jest (GI4).

Dla warunku (GI2), używamy twierdzenia Stone'a-Weierstrassa: D jest algebrą funkcji ciągłych na zbiorze zwartym, która rozdziela punkty. Istotnie, dla $\eta \neq \zeta$ istnieje $x \in S$ takie, że $\eta(x) \neq \zeta(x)$. Funkcja $f(\xi) = \xi(x)$ rozdziela η i ζ . Algebra D zawiera funkcje stałe, więc $\overline{D} = C(\{0, 1\}^V)$.

Dla (GI3), załóżmy $f \in D$, $\lambda \geq 0$, oraz $f - \lambda Lf = g$. Ponieważ $\{0, 1\}^V$ jest zbiorem zwartym i f jest ciągła, istnieje takie η , dla którego f osiąga swoje minimum. Wtedy $Lf(\eta) \geq 0$, więc

$$\min_{\zeta} f(\zeta) = f(\eta) \geq g(\eta) \geq \min_{\zeta} g(\zeta).$$

Warunek (GI5) wynika z faktu, że $1 \in D$ oraz $L1 = 0$.

wykład 9: Konstrukcja systemów spinowych

Piotr Dyszewski

28-11-24

Dla $x \in V$ oraz $\eta \in \{0, 1\}^V$ definiujemy $\eta^{(x)} \in \{0, 1\}^V$ wzorem

$$\eta^{(x)}(y) = \begin{cases} \eta(y), & y \neq x \\ 1 - \eta(x), & y = x \end{cases}.$$

Dla f pochodzącego z odpowiedniego podzbioru $C_0(\{0, 1\}^V)$ chcemy położyć

$$Lf(\eta) = \sum_{x \in V} c(x, \eta) [f(\eta^{(x)}) - f(\eta)].$$

Okazuje się, że dokładne napisanie dziedziny jest problematyczne. Aby obejść tę trudność rozważmy

$$D = \left\{ f \in C(\{0, 1\}^V) : \|f\|_o := \sup_{\eta} \sum_x |f(\eta^{(x)}) - f(\eta)| < \infty \right\}.$$

Konstrukcja systemów spinowych cd.

Aby sprawdzić (GI4) musimy wyprowadzić ograniczenie dla rozwiązań równania $f - \lambda Lf = g$. Niech

$$\epsilon = \inf_{u, \eta} [c(u, \eta) + c(u, \eta_u)] \quad \text{oraz} \quad \gamma(x, u) = \sup_{\eta} |c(x, \eta_u) - c(x, \eta)|.$$

Zauważmy, że $\gamma(x, u)$ mierzy stopień, w jakim intensywność zmiany w miejscu x zależy od konfiguracji w miejscu u . Niech $\ell_1(V)$ będzie przestrzenią Banacha funkcji $\alpha : V \rightarrow \mathbb{R}$, które spełniają

$$\|\alpha\| := \sum_x |\alpha(x)| < \infty.$$

Macierz γ definiuje operator Γ na $\ell_1(S)$ przez

$$\Gamma\alpha(u) = \sum_{x: x \neq u} \alpha(x) \gamma(x, u).$$

Operator ten jest dobrze zdefiniowany i ograniczony, pod warunkiem że

$$M := \sup_x \sum_{u: u \neq x} \gamma(x, u) < \infty,$$

a wtedy $\|\Gamma\| = M$.

Dla $f \in C(\{0, 1\}^V)$ i $x \in S$, niech

$$\Delta f(x) = \sup_{\eta} \left| f\left(\eta^{(x)}\right) - f(\eta) \right|.$$

Wtedy $\|f\|_o = \|\Delta f\|_{l_1(V)}$. Oto oszacowanie, którego potrzebujemy.

Fakt 0.2. *Założmy, że spełniony jest jeden z warunków*

- $f \in D$,
- f jest ciągła i

$$c(x, \cdot) \equiv 0 \text{ dla wszystkich oprócz skończonej liczby } x \in V. \quad (27)$$

Wówczas jeśli $f - \lambda Lf = g \in D$, $\lambda > 0$, oraz $\lambda M < 1 + \lambda\epsilon$, to

$$\Delta f \leq [(1 + \lambda\epsilon)I - \lambda\Gamma]^{-1} \Delta g, \quad (28)$$

gdzie nierówność zachodzi współrzędna po współrzędnej, a odwrotność jest zdefiniowana przez nieskończony szereg

$$[(1 + \lambda\epsilon)I - \lambda\Gamma]^{-1} \alpha = \frac{1}{1 + \lambda\epsilon} \sum_{k=0}^{\infty} \left(\frac{\lambda}{1 + \lambda\epsilon} \right)^k \Gamma^k \alpha. \quad (29)$$

Proof. Zauważmy, że szereg w (29) jest zbieżny dla $\alpha \in \ell_1(V)$ na mocy założenia $\lambda M < 1 + \lambda\epsilon$. Pisząc $f - \lambda Lf = g$ w punktach η oraz $\eta^{(u)}$, odejmując i zauważając że $(\eta^{(u)})^{(u)} = \eta$, otrzymujemy

$$[f(\eta^{(u)}) - f(\eta)][1 + \lambda c(u, \eta) + \lambda c(u, \eta^{(u)})] = [g(\eta^{(u)}) - g(\eta)] + \lambda \sum_{x: x \neq u} \left\{ c(x, \eta^{(u)})[f((\eta^{(u)})^{(x)}) - f(\eta^{(u)})] - c(x, \eta)[f(\eta^{(x)}) - f(\eta)] \right\}. \quad (30)$$

Ponieważ wartości $f(\eta^{(u)}) - f(\eta)$, gdy η zmienia się a u jest ustalone, tworzą zbiór symetryczny, a ta różnica jest funkcją ciągłą η , dla każdego u istnieje takie η , że

$$f(\eta^{(u)}) - f(\eta) = \sup_{\zeta} |f(\zeta^{(u)}) - f(\zeta)| = \Delta f(u).$$

Stąd,

$$f(\zeta^{(u)}) - f(\zeta) \leq f(\eta^{(u)}) - f(\eta)$$

dla każdej ζ . Stosując to dla $\zeta = \eta^{(x)}$ i przekształcając, otrzymujemy

$$f((\eta^{(u)})^{(x)}) - f(\eta^{(u)}) = f((\eta^{(x)})^{(u)}) - f(\eta^{(u)}) \leq f(\eta^{(x)}) - f(\eta),$$

Używając tej nierówności w (30),

$$\begin{aligned} \Delta f(u)(1 + \lambda\epsilon) &\leq \Delta f(u)[1 + \lambda c(u, \eta) + \lambda c(u, \eta^{(u)})] \\ &\leq \Delta g(u) + \lambda \sum_{x: x \neq u} [c(x, \eta^{(u)}) - c(x, \eta)] [f(\eta^{(x)}) - f(\eta)] \\ &\leq \Delta g(u) + \lambda \sum_{x: x \neq u} \gamma(x, u) \Delta f(x). \end{aligned} \quad (31)$$

Jeśli (27) zachodzi, to tylko skończona liczba wyrazów po prawej stronie jest niezerowa, więc przy któregośkolwiek z założeń faktu $\Gamma \Delta_f$ jest dobrze określona. Dlatego (31) można zapisać jako

$$(1 + \lambda\epsilon) \Delta f \leq \Delta g + \lambda \Gamma \Delta f.$$

Iteracja tej nierówności prowadzi to do

$$\Delta f \leq \frac{1}{1 + \lambda\epsilon} \sum_{k=0}^n \left(\frac{\lambda}{1 + \lambda\epsilon} \right)^k \Gamma^k \Delta g + \left(\frac{\lambda}{1 + \lambda\epsilon} \right)^{n+1} \Gamma^{n+1} \Delta f.$$

Jeżeli rozważymy teraz $n \rightarrow \infty$, dostaniemy (28). □

Twierdzenie 0.17. Załóżmy, że $M < \infty$. Wtedy \bar{L} jest generatorem infinitesimalnym półgrupy Fellera $T = (T(t))_{t \in \mathbb{R}_+}$. Ponadto,

$$\Delta T(t)f \leq e^{-t\epsilon} e^{t\Gamma} \Delta f. \quad (32)$$

W szczególności, jeśli $f \in D$, to $T_t f \in D$ oraz

$$\|T(t)f\|_o \leq e^{(M-\epsilon)t} \|f\|_o. \quad (33)$$

Proof. Własności (GI1), (GI2), (GI3) i (GI5) z Definicji ?? zachodzą dla (L, D) są i są dziedziczone przez \bar{L} z Faktu 0.1. Aby sprawdzić warunek (GI4) weźmy wstępujący ciąg $V_n \subseteq V$ taki, że $\bigcup_n V_n = V$. Niech

$$L_n f(\eta) = \sum_{x \in V_n} c(x, \eta) [f(\eta^{(x)}) - f(\eta)], \quad f \in C(\{0, 1\}^V). \quad (\#eq : 4.10)$$

To jest generator dla systemu spinowego, w którym współrzędne

$$(\eta(x) : x \notin V_n)$$

są stałe w czasie. Ponieważ L_n jest ograniczonym generatorem, spełnia

$$\mathcal{R}(I - \lambda L_n) = C(\{0, 1\}^V)$$

dla dostatecznie małych $\lambda > 0$. Dla $g \in D$, możemy zdefiniować $f_n \in C(\{0, 1\}^V)$ przez $f_n - \lambda L_n f_n = g$. Ponieważ L_n spełnia (27), jeśli λ jest wystarczająco małe, tak że $\lambda M < 1 + \lambda \epsilon$, wtedy $f_n \in D$ zgodnie z Faktem 0.2. W związku z tym możemy położyć

$$g_n = f_n - \lambda L f_n \in \mathcal{R}(I - \lambda L).$$

Niech $K = \sup_{x, \eta} c(x, \eta) < \infty$, wtedy z Faktu 0.2,

$$\begin{aligned} \|g_n - g\| &= \lambda \|(L - L_n)f_n\| \leq \lambda K \sum_{x \notin V_n} \Delta f_n(x) \\ &\leq \lambda K \sum_{x \notin V_n} [(1 + \lambda \epsilon)I - \lambda \Gamma]^{-1} \Delta g(x). \end{aligned} \quad (34)$$

Ponieważ $\Delta g \in \ell_1(V)$, prawa strona (34) dąży do zera, gdy $n \rightarrow \infty$, więc $g_n \rightarrow g$. Stąd $g \in \text{cl}(\mathcal{R}(I - \lambda L))$, więc wnioskujemy, że $D \subseteq \text{cl}(\mathcal{R}(I - \lambda L))$. Ponieważ D jest gęste w $C(\{0, 1\}^V)$, widzimy, że $\mathcal{R}(I - \lambda L)$ jest również

gęste. Zgodnie z Faktem 0.1, $\mathcal{R}(I - \lambda\bar{L})$ musi być domkniętym podzbiorem $C(\{0, 1\}^V)$. Zatem

$$\mathcal{R}(I - \lambda\bar{L}) = C(\{0, 1\}^V)$$

To kończy weryfikację, że \bar{L} jest generatorem infinitezymalnym.

Przechodząc do drugiego stwierdzenia, zapiszmy (28) jako

$$\Delta_{(I - \lambda L)^{-1}} g \leq [(1 + \lambda\epsilon)I - \lambda\Gamma]^{-1} \Delta g,$$

a następnie iterujemy, aby uzyskać

$$\Delta_{(I - \frac{t}{n}L)^{-1}} g \leq \left[\left(1 + \frac{t}{n}\epsilon\right) I - \frac{t}{n}\Gamma \right]^{-n} \Delta g.$$

Przechodząc do granicy otrzymujemy (32). □

Wykład 10: Procesy dualne

Piotr Dyszewski

12-12-2014

Procesy dualne

Często zdarza się, że wartość oczekiwaną związaną z jednym skomplikowany procesem jesteśmy w stanie wyrazić jako wartość oczekiwaną związaną z innym procesem, który jest o wiele postrzy. Tego typu relacje pozwalają na przydatne reprezentacje wielkości występujących w procesach Fellera.

Definicja 1. Niech $X_1 = (X_1(t))_{t \in \mathbb{R}}$ i $X_2 = (X_2(t))_{t \in \mathbb{R}}$ będą procesami Fellera odpowiednio na przestrzeniach S_1 i S_2 . Dla mierzalnej i ograniczonej funkcji H na $S_1 \times S_2$, procesy te są nazywane dualnymi względem H , jeśli

$$\mathbf{E}_{x_1}[H(X_1(t), x_2)] = \mathbf{E}_{x_2}[H(x_1, X_2(t))]$$

dla każdego $t \geq 0$ oraz $x_i \in S_i$.

Powyższe pojęcie w zupełnej ogólności jest problematyczne. Naturalnym jest oczekiwać, że () mówi coś o relacji między X_1 oraz X_2 . Zauważmy, że każde dwa procesy są dualne względem funkcji stałej. Jednakże charakter relacji między X_1 a X_2 zależy bardzo mocno od konteksty i podyktowany jest przez funkcję H .

Niech X_a oraz X_r będą ruchami Browna na $S_1 = S_2 = [0, +\infty)$ odpowiednio zabitymi i odbitymi w zerze. Procesy te są dualne względem funkcji

$$H(x, y) = \mathbf{1}_{\{x \leq y\}}.$$

Wówczas relacja () zapisuje się jako

$$\mathbf{P}_{x_1}[X_1(t) \leq x_2] = \mathbf{P}_{x_2}[x_1 \leq X_2(t)].$$

W przypadku wspomnianych wersji ruchu Browna

$$\mathbf{P}_x[X_a(t) \leq y] = \mathbf{P}_y[x \leq X_r(t)].$$

Oba prawdopodobieństwa są równe

$$\mathbb{P}[B_t \geq x - y] + \mathbb{P}[B_t \geq x + y],$$

gdzie $B = (B_t)_{t \in \mathbb{R}_+}$ jest standardowym ruchem Browna na \mathbb{R} . Dokładne sprawdzenie wspomnianej równości pozostawiamy jako zadanie.

Założmy teraz, że H jest ciągła. W świetle definicji procesy Fellera X_1 oraz X_2 z półgrupami Fellera odpowiednio $T_1 = (T_1(t))_{t \in \mathbb{R}_+}$ oraz $T_2 = (T_2(t))_{t \in \mathbb{R}_+}$ są H -dualne wtedy i tylko wtedy, gdy

$$T_1(t)H(\cdot, s_2)(s_1) = T_2(t)H(s_1, \cdot)(s_2)$$

dla wszystkich $t \geq 0$, $s_1 \in S_1$ oraz $s_2 \in S_2$. W przypadku, gdy definiujemy procesy Fellera przez ich opis infinitesimalny wygodniejsze jest kryterium wyrażone w terminach generatorów.

Twierdzenie 3. Niech X_1 i X_2 będą generowane odpowiednio przez $(L_1, \mathcal{D}(L_1))$ oraz $(L_2, \mathcal{D}(L_2))$, Założmy, że dla każdych $s_1 \in S_1$ oraz $s_2 \in S_2$

$$H(\cdot, s_2) \in \mathcal{D}(L_1) \quad \text{oraz} \quad H(s_1, \cdot) \in \mathcal{D}(L_2).$$

Jeżeli dodatkowo

$$L_1H(\cdot, s_2)(s_1) = L_2H(s_1, \cdot)(s_2)$$

dla wszystkich $x_1 \in S_1$ oraz $x_2 \in S_2$. Wówczas X_1 i X_2 są dualne względem H .

Proof. Rozumowanie przeprowadzimy jedynie w przypadku przeliczalnej S_2 i ograniczonego L_2 . Wówczas X_2 jest łańcuchem Markowa z q -macierzą $q = (q(x, y))_{x, y \in S_2}$. Przypomnijmy, że wówczas

$$T_2(t)f(s_2) = \mathbf{E}_{s_2}[f(X_2(t))] = \sum_{y \in S_2} \mathbf{P}_{s_2}[X_2(t) = y]f(y).$$

Generator L_2 zadany jest wówczas przez

$$L_2 f(s_2) = \left. \frac{d}{dt} \right|_{t=0} T_2(t) f(s_2) = \sum_{y \in S_2} \left. \frac{d}{dt} \right|_{t=0} \mathbf{P}_{s_2}[X_2(t) = y] f(y) = \sum_{y \in S_2} q(s_2, y) f(y).$$

Rozważmy

$$u(t, x_1, x_2) = \mathbf{E}_{x_1} H(X_1(t), x_2) = T_1(t) H(\cdot, x_2)(x_1)$$

Na mocy Twierdzenia 0.10,

$$\begin{aligned} \frac{d}{dt} u(t, x_1, x_2) &= T_1(t) L_1 H(\cdot, x_2)(x_1) = T_1(t) L_2 H(x_1, \cdot)(x_2) \\ &= \sum_{y \in S_2} q(x_2, y) T_1(t) H(\cdot, y)(x_1) = \sum_y q(x_2, y) u(t, x_1, y) = L_2 u(t, x_1, \cdot)(x_2). \end{aligned}$$

Dodatkowo $u(0, x_1, x_2) = H(x_1, x_2)$. Z drugiej strony funkcja

$$v(t, x_1, x_2) = T_2(t) H(x_1, \cdot)(x_2)$$

również spełnia $v(0, x_1, x_2) = H(x_1, x_2)$. Wystarczy zatem uzasadnić jedyność tego zagadnienia. Rozważmy w tym celu $h = v - u$. Wówczas

$$h(t, x_1, x_2) = \int_0^t L_2 h(s, x_1, \cdot) ds$$

Niech

$$h^*(t, x_1) = \sup_{s \leq t, x_2 \in S_2} h(s, x_1, x_2).$$

Wówczas

$$h^*(t, x_1) \leq \int_0^t \|L_2\| h^*(s, x_1) ds.$$

Powyższa nierówność związa się do

$$\frac{d}{dt} \left(e^{-\|L_2\|t} \int_0^t h^*(s, x_1) ds \right) \leq 0.$$

Co po całkowaniu daje

$$e^{-\|L_2\|t} \int_0^t h^*(s, x_1) ds \leq 0.$$

Skoro lewa strona jest niewątpliwie nieujemna, to $h \equiv 0$ za co za tym idzie $u \equiv v$. Ostatnie równość jest równoważna z dowodzoną tezą. \square

Przykład 4. Niech X_1 i X_2 będą spacerami losowymi na \mathbb{Z} z q -macierzami odpowiednio

$$q_1(x, x+1) = \beta, \quad q_1(x, x-1) = \delta$$

oraz

$$q_2(x, x+1) = \delta, \quad q_2(x, x-1) = \beta$$

The Voter model

Niech V będzie przeliczalnym zbiorem z topologia dyskretną. Chcemy modelować proces rozwoju opinii wśród osobników reprezentowanych przez elementy V . Zakładać będziemy, że w każdej chwili czasu $t \geq 0$ każdy osobnik $x \in V$ reprezentuje jedną z dwóch opinii $\eta_t(x) \in \{0, 1\}$ na zadany temat. Założymy, że dane są nieujemne liczby $q(x, y)$ dla $x \neq y$. Wielkość $q(x, y)$ będzie intensywnością z jaką x przejmuje opinię y o ile oba osobniki reprezentują różne opinie. Zakładać będziemy, że

$$M = \sup_{x \in V} \sum_{u: u \neq x} q(x, u) < \infty.$$

Model głosowania (the Voter model) η_t to system spinowy z

$$c(x, \eta) = \sum_{y: \eta(y) \neq \eta(x)} q(x, y).$$

Innymi słowy jest to proces Fellera generowany przez

$$Lf(\eta) = \sum_{x \in V} c(x, \eta) \left(f(\eta^{(x)}) - f(\eta) \right).$$

Techniczne szczegóły związane z dziedziną L zostały przedyskutowane w poprzednim rozdziale. Najważniejsze jest, że z Twierdzenia @ref(thm:4.3) wiemy, że proces ten jest dobrze określony (domknięcie L jest generatorem infinitezymalnym).

Przykład 5. Założymy, że V jest wyposażone w strukturę grafu o ograniczonym stopniu. Chcemy modelować przypadek w którym każdy z wierzchołków $x \in V$ może wchodzić w interakcję jedynie ze swoimi bezpośrednimi sąsiadami (i to od nich zapożycza opinie). Rozważmy $q(x, y) = \mathbf{1}_{x \sim y}$. Wówczas

$$M = \sup_{x \in V} \sum_{u: u \neq x} q(x, u) = \sup_{x \in V} \deg(x) < \infty.$$

Zauważmy, że model głosowania posiada dwa stany stacjonarne $\eta \equiv 1$ oraz $\eta \equiv 0$. Naszym głównym celem jest sprawdzenie, czy istnieją inne (nietrywialne) rozkłady stacjonarne.

Aby tego dokonać posłużymy się procesem dualnym do $(\eta_t)_t$. Ustalmy $t \geq 0$ i $x \in V$. Skoro przy zmianach opinia w x jest zapożyczana od innych osobników, chcąc zbadać wartość $\eta_t(x)$ rozważmy t_1 -moment ostatniej zmiany opinii przez x , czyli

$$t_1 = \sup_{s \leq t} \{\eta_{s-}(x) \neq \eta_s(x)\}$$

Jeżeli zbiór czasów pod kresem górnym jest pusty, to x nie zmienił zdania na odcinku czasu $[0, t]$, więc $\eta_t(x) = \eta_0(x)$. W chwili t_1 , x przyjął tę samą opinię co pewien x_1 (co się dzieje z intensywnością $q(x, x_1)$), czyli $\eta_t(x) = \eta_{t_1}(x_1)$. Chcąc ustalić wartość $\eta_{t_1}(x_1)$ rozważamy ostatni moment, w którym x_1 zmienił opinię

$$t_2 = \sup_{s \leq t_1} \{\eta_{s-}(x_1) \neq \eta_s(x_1)\}$$

Jeżeli zbiór pod kresem górnym jest pusty, to x_1 na przedziale czasowym $[0, t_1]$ nie zmienił zdania i $\eta_t(x) = \eta_{t_1}(x_1) = \eta_0(x_1)$. Postępując iteracyjne dostajemy ciąg czasów $t \geq t_1 > t_2 > \dots > t_N$ taki, że

$$\eta_t(x) = \eta_{t_1}(x_1) = \dots = \eta_{t_N}(x_N) = \eta_0(x_N).$$

Przy czym przejście z x_k do x_{k+1} dzieje się z intensywnością $q(x_k, x_{k+1})$. Skonstruowana w ten sposób ścieżka (x, x_1, \dots, x_N) ma taki sam rozkład jak ścieżka łańcucha Markowa $Y_x = (Y_x(t))_{t \in \mathbb{R}_+}$ z q -macierzą $(q(x, y))_{x, y \in V}$. Oznacza to, że

$$\eta_t(x) \stackrel{d}{=} \eta_0(Y_x(t)).$$

Chcąc zbadać teraz rozkład łączny $(\eta_t(x), \eta_t(y))$ dla $x, y \in V$ możemy wykonać podobną konstrukcję na podstawie zmian opinii któregośkolwiek z elementów pary. Dostaniemy w ten sposób ciąg czasów $t \geq s_1 > s_2 > \dots > s_M \geq 0$ taki, że

$$\eta_t(x) = \eta_{s_1}(x_1) = \dots = \eta_{s_M}(x_M) = \eta_0(x_M).$$

oraz

$$\eta_t(y) = \eta_{s_1}(y_1) = \dots = \eta_{s_M}(y_M) = \eta_0(y_M).$$

Kluczowa jest następująca własność. Jeżeli $x_j = y_j$ dla pewnego $j \geq 1$, to $x_k = y_k$ dla wszystkich $k \geq j$. Istotnie, jeżeli j jest najmniejszą taką liczbą,

że $x_j = y_j$, to oznacza to że x_{j-1} przejął opinię x_j . Po czym zanim x_j zmienił opinię, to y_{j-1} przejął opinię x_j . Oznacza to, że

$$(\eta_t(x), \eta_y(y)) \stackrel{d}{=} (\eta_0(Y_x(t)), \eta_0(Y_y(t))),$$

gdzie Y_x i Y_y są łańcuchami Markowa na V z zadaną q -macierzą takimi, że jeżeli w pewnym momencie się spotkają, to od tego momentu zaczynają się poruszać się razem. Podobny komentarz możemy napisać dla wektora $\eta_x(t)$ dowolnej długości: jego rozkład będziemy mogli wyrazić przez kolekcję spacerów losowych na V , które się zlewają w momencie spotkania.

Powyżej znajduje się symulacja zlewających się spacerów losowych na torusie dyskretnym. Cząsteczki poruszają się według symetrycznego spaceru losowego.

Aby wprowadzić ten proces bardziej formalnie, rozważmy $S_2(N)$ - zbiór wszystkich $A \subseteq V$ o liczebności nie większej niż N . Rozważmy $\{Q(A, B)\}_{A, B \in S_2}$ dane przez

$$Q(A, (A \setminus \{x\}) \cup \{y\}) = q(x, y), \quad x \in A, y \notin A;$$

oraz

$$Q(A, A \setminus \{x\}) = \sum_{y \in A, y \neq x} q(x, y), \quad x \in A.$$

Taki wybór instancyjności przejść odpowiada dokładnie zlewającym się spacerom losowym. Generator takiego procesu jest ograniczony z normą

$$\sum_{B \neq A} Q(A, B) = \sum_{x \in A} \sum_{y \neq x} q(x, y) \leq MN.$$

Pokażemy, że zlewające się spacery losowe $A = (A_t)_{t \in \mathbb{R}}$ (proces o intensywnościach danych powyżej) jest dualny do modelu głosowania $(\eta_t)_{t \in \mathbb{R}}$ z funkcją

$$H(\eta, A) = \prod_{x \in A} \eta(x) = 1_{\{\eta=1 \text{ on } A\}},$$

Trajektorie $|A_t|$ są nierosnące, co czyni go bardzo użytecznym w badaniu modelu głosowania,

Twierdzenie 6. Procesy $(\eta_t)_t$ i $(A_t)_t$ są dualne względem $H(\eta, A)$.

Proof. Sprawdzimy, że zachodzą założenia Twierdzenia 3. Niech L będzie generatorem modelu głosowania. Ponieważ $H(\eta, A)$ zależy od η tylko poprzez $\{\eta(x), x \in A\}$,

$$\begin{aligned}
LH(\cdot, A)(\eta) &= \sum_{\substack{x \in A, y \in S \\ \eta(y) \neq \eta(x)}} q(x, y)[H(\eta_x, A) - H(\eta, A)] \\
&= \sum_{\substack{x \in A, y \in S \\ \eta(y) \neq \eta(x)}} q(x, y)[1 - 2\eta(x)]H(\eta, A \setminus \{x\}) \\
&= \sum_{x \in A, y \in S} q(x, y)H(\eta, A \setminus \{x\})[\eta(y) - \eta(x)] \\
&= \sum_{x \in A, y \in S} q(x, y)[H(\eta, (A \setminus \{x\}) \cup \{y\}) - H(\eta, A)] \\
&= \sum_B q(A, B)[H(\eta, B) - H(\eta, A)].
\end{aligned}$$

□

Wykład 11: The voter model: przypadek rekurencyjny

2024-12-19

Piotr Dyszewski

Zobaczmy teraz korzyści płynące z udowodnionej właśnie dualności. Jeżeli $(\eta_t)_{t \in \mathbb{R}_+}$ jest układem spinowym, to dla $A \subseteq V$ oraz miary probabilistycznej μ na $\{0, 1\}$ prawdopodobieństwo

$$\mathbf{P}_\mu[\eta_t \equiv 1 \text{ na } A]$$

zależy od μ w skomplikowany sposób. W przypadku modelu głosowania, możemy wykorzystać proces dualny do wyznaczenia wygodnej reprezentacji

$$\begin{aligned} \mathbf{P}_\mu[\eta_t \equiv 1 \text{ na } A] &= \int \mathbf{E}_\eta H(\eta_t, A) \mu(d\eta) = \int \mathbf{E}_A H(\eta, A_t) \mu(dA) \\ &= \int \left[\sum_B \mathbf{P}_A[A_t = B] H(\eta, B) \right] \mu(d\eta) = \sum_B \mathbf{P}_A[A_t = B] \mu\{\eta : \eta \equiv 1 \text{ na } B\}. \end{aligned}$$

Dlatego rozkład łączny $(\eta_t(x))_{x \in A}$ zależy tylko od rozkładów $|A|$ współrzędnych względem rozkładu początkowego μ . Załóżmy teraz, że początkowy rozkład dla modelu głosowania to miara produktowa ν_ρ dla $\rho \in [0, 1]$ względem której wszystkie współrzędne są niezależne i mają rozkład $(1 - \rho)\delta_0 + \rho\delta_1$. Wówczas

$$\nu_\rho\{\eta : \eta \equiv 1 \text{ na } A\} = \rho^{|A|}.$$

Mamy

$$\mathbf{P}_{\nu_\rho}[\eta_t(x) = 1] = \mathbf{P}_{\nu_\rho}[\eta_t \equiv 1 \text{ na } \{x\}] = \sum_{y \in V} \mathbf{P}_{\{x\}}[A_t = \{y\}] \rho = \rho$$

oraz

$$\begin{aligned}\mathbf{P}_{\nu_\rho}[\eta_t(x) = \eta_t(y) = 1] &= \mathbf{P}_{\nu_\rho}(\eta_t \equiv 1 \text{ na } \{x, y\}) \\ &= \rho \mathbf{P}_{\{x, y\}}[|A_t| = 1] + \rho^2 \mathbf{P}_{\{x, y\}}[|A_t| = 2].\end{aligned}$$

Opinie $\eta_t(x)$ i $\eta_t(y)$ stają się zgodne w granicy $t \rightarrow \infty$ wtedy i tylko wtedy, gdy dwie niezależne kopie łańcucha spotykają się z prawdopodobieństwem 1. Tę ogólną zasadę zilustrujemy na przykładzie $V = \mathbb{Z}^d$.

The Voter model na \mathbb{Z}^d

Naszym głównym zastosowaniem dualności będzie znalezienie wszystkich rozkładów stacjonarnych dla modelu głosowania. Podobna analiza może być użyta do określenia jego asymptotycznego zachowania, gdy $t \rightarrow \infty$. Dla uproszczenia, zakładamy od tej pory, że $S = \mathbb{Z}^d$. Zakładać też będziemy, że kroki stowarzyszonego łańcucha są niezmiennicze na przesunięcia, czyli przejścia z $x_1 \in \mathbb{Z}^d$ oraz przejścia z $x_2 \in \mathbb{Z}^d$ dzieją się z takimi samymi intensywnościami. Wówczas $q(x, y) = q(0, y - x)$. Niech $X_i = (X_i(t))_{t \in \mathbb{R}_+}$ dla $i \in \mathbb{N}$ będą niezależnymi błędzeniami losowymi z intensywnościami przejść $q(x, y)$ i punktami początkowymi x_1, x_2, \dots . Analiza dzieli się na dwa przypadki, w zależności od tego, czy symetryzowane błędzenie losowe $Z(t) = X_1(t) - X_2(t)$ jest rekurencyjne czy przejściowe, ponieważ odpowiada to sytuacji, czy $X_1(t)$ i $X_2(t)$ spotykają się z prawdopodobieństwem jeden.

Przypadek rekurencyjny

Zacniemy od przypadku rekurencyjnego, w którym model głosowania jest prostszy w analizie. Niech δ_0 i δ_1 będą miarami probabilistycznymi na układach na \mathbb{Z}^d skupionymi na odpowiednio $\eta \equiv 0$ oraz $\eta \equiv 1$.

Twierdzenie 0.18. *Załóżmy, że symetryzowane błędzenie losowe Z jest rekurencyjne.*

- Dla każdego η oraz $x_1, x_2 \in \mathbb{Z}^d$,

$$\lim_{t \rightarrow \infty} \mathbf{P}_\eta[\eta_t(x_1) \neq \eta_t(x_2)] = 0.$$

- Każda miara stacjonarna dla $(\eta_t)_{t \in \mathbb{R}_+}$ jest postaci $\lambda \delta_1 + (1 - \lambda) \delta_0$ dla pewnej $\lambda \in [0, 1]$.

- Dla miary probabilistycznej μ oraz $\lambda \in [0, 1]$,

$$\lim_{t \rightarrow \infty} \mathbf{P}_\mu[(\eta_t(x_j))_{j \leq n} = (\epsilon_j)_{j \leq n}] = \lambda \mathbf{1}_{\{\epsilon_1 = \dots = \epsilon_n = 1\}} + (1 - \lambda) \mathbf{1}_{\{\epsilon_1 = \dots = \epsilon_n = 0\}} \quad (35)$$

dla każdych x_j i ϵ_j , wtedy i tylko wtedy, gdy

$$\lim_{t \rightarrow \infty} \sum_y p_t(x, y) \mu\{\eta : \eta(y) = 1\} = \lambda \text{ dla każdego } x \in \mathbb{Z}^d, \quad (36)$$

gdzie $p_t(x, y)$ są prawdopodobieństwami przejścia dla $X_i(t)$.

Punkt (a) mówi, że dla dowolnego skończonego $A \subseteq \mathbb{Z}^d$ opinie $(\eta_t(x))_{x \in A}$ stają się zgodne z bardzo dużym prawdopodobieństwem. Oznacza to, że jedyne możliwe układy graniczne są stałe (punkt (b)). Prawdopodobieństwa wystąpienia w granicy poszczególnych układów stałych są scharakteryzowane w punkcie (c). W szczególności pierwsza granica w punkcie (c) istnieje wtedy i tylko wtedy, gdy istnieje druga. Warunek (35) jest równoważny

$$\lim_{t \rightarrow \infty} \mathbf{P}_\mu[\eta_t \in \cdot] = \lambda \delta_1(\cdot) + (1 - \lambda) \delta_0(\cdot)$$

w sensie słabej zbieżności miar. Dokładniej dla każdej $f \in C(\{0, 1\}^{\mathbb{Z}^d})$,

$$\lim_{t \rightarrow \infty} \mathbf{E}_\mu[f(\eta_t)] = \lambda f(\mathbf{1}) + (1 - \lambda) f(\mathbf{0}).$$

Dzieje się tak, ponieważ topologia słabej zbieżności na przeliczalnym produkcie jest scharakteryzowana przez zbieżność rozkładów skończone wymiarowych.

Proof. Aby uzasadnić punkt (a) rozważmy

$$\tau = \inf\{t > 0 : X_1(t) = X_2(t)\} = \inf\{t > 0 : Z(t) = 0\}.$$

Wówczas dla dowolnej konfiguracji początkowej η ,

$$\mathbf{P}_\eta[\eta_t(x_1) \neq \eta_t(x_2)] = \mathbb{P}[\eta(X_1(t)) \neq \eta(X_2(t)), \tau > t] \leq \mathbb{P}[\tau > t],$$

co zbiega do zera, gdy $t \rightarrow \infty$, z założenia rekurencyjności Z .

Dla części (b), weźmy π stacjonarną dla $(\eta_t)_{t \in \mathbb{R}}$. Wtedy

$$\pi[\eta : \eta(x_1) \neq \eta(x_2)] = \mathbf{P}_\pi[\eta(x_1) \neq \eta(x_2)] = \int \mathbf{P}_\eta[\eta_t(x_1) \neq \eta_t(x_2)] \mu(d\eta),$$

co zbiega do zera, gdy $t \rightarrow \infty$, z części (a). Dlatego η jest stała π -prawie wszędzie. Dla dowodu części (c) weźmy dowolny rozkład początkowy μ . Chcąc pokazać, że rozkład graniczny jest stały wystarczy porównać go z dowolnym ustalonym punktem. Dla $x_1 \in \mathbb{Z}^d$,

$$\begin{aligned} \mathbf{P}_\mu[\eta_t(x_1) = 1] &= \int \mathbf{P}_\eta[\eta_t(x_1) = 1] \mu(d\eta) = \\ &= \int \mathbb{P}[\eta(X_1(t)) = 1] \mu(d\eta) = \sum_y p_t(x_1, y) \mu(\eta : \eta(y) = 1). \end{aligned} \quad (37)$$

Dla dowolnego skończonego $A \subseteq \mathbb{Z}^d$ takiego, że $x_1 \in A$,

$$\mathbf{P}_\mu[\eta_t(x_1) = 1] - \mathbf{P}_\mu[\eta_t \equiv 1 \text{ na } A] = \int \mathbf{P}_\eta[\eta_t(x_1) = 1] - \mathbf{P}_\eta[\eta_t \equiv 1 \text{ na } A] \mu(d\eta).$$

Co z części (a) zbiega do zera. Mamy zatem

$$\sum_y p_t(x_1, y) \mu(\eta : \eta(y) = 1) = o(1) + \mathbf{P}_\mu[\eta_t \equiv 1 \text{ na } A].$$

Stąd (35) oraz (36) istotnie są równoważne. \square

Remark. Biorąc pod uwagę stwierdzenie z części (a), można przypuszczać, że $\eta_t(x)$ zmienia wartości tylko skończoną liczbę razy dla każdego x . Nie zawsze jest to prawdą. Załóżmy na przykład, że $d = 1$, $q(x, y) = 1$ jeśli $|x - y| = 1$, a $q(x, y) = 0$ jeśli $|x - y| > 1$. Jeśli początkowa konfiguracja ma formę $\dots 111000\dots$, to konfiguracja w czasie t ma ten sam typ, z dokładnością do lokalizacji najbardziej wysuniętej na prawo jedynek, która porusza się zgodnie z prostym symetrycznym spacerem losowym na \mathbb{Z} . Ponieważ to błędzenie losowe jest rekurencyjne, $\eta_t(x)$ zmienia wartości nieskończoną liczbę razy.

Symulacja modelu głosowania w jednym wymiarze. Warunek początkowy jest postaci 11110000. W momencie, w którym cały zakres staje się monochromatyczny jest on automatycznie rozszerzany na prawo o 0000 i lewo o 111111.