Metody probabilistyczne i statystyka - wykład piętnasty[1]
Elementy procesów stochastycznych – łańcuchy Markowa

dr Jarosław Kotowicz

Instytut Informatyki Uniwersytet w Białymstoku

wersja z roku ak.2020/21

Spis treści

Podstawowe pojęcia. I

Niech (Ω, Σ, P) będzie przestrzenią probabilistyczną, S zbiorem co najwyżej przeliczalnym zbiorem, zaś $(X_n)_{n \geqslant 0}$ ciągiem zmiennych losowych określonych na przestrzeni (Ω, Σ, P) o wartościach w S.

Definicja 1

Ciąg zmiennych losowych $(X_n)_{n \geqslant 0}$ nazywamy łańcuchem Markowa wtedy i tylko wtedy, gdy dla dowolnego $n \in \mathbb{N}$ i dowolnych $\{s_0, \ldots, s_n\} \subset S$ zachodzi

$$P(\{X_n = s_n\} | \{X_k = s_k, k \in \overline{0, n-1}\}) = \\ P(\{X_n = s_n\} | \{X_{n-1} = s_{n-1}\}), \tag{1}$$

o ile $P(\{X_k = s_k \wedge k = 0, \ldots n-1\}) > 0$.

Podstawowe pojęcia. II

Uwaga 1

Elementy zbioru S nazywamy stanami. Ponieważ zbiór S jest co najwyżej przeliczalny więc jego elementy można ustawić w ciąg (tzn. zadać porządek liniowy) numerując kolejnymi liczbami naturalnymi i zerem. W związku z tym stany s_k będziemy utożsamiać (oznaczać) z jego indeksem.

Definicja 2

Macierz $\mathcal{P} = [p_{ij}]_{i,j\in S}$ nazywamy macierzą przejścia na S (macierzą stochastyczną) wtedy i tylko wtedy, gdy

$$\forall_{i,j\in S}\, p_{ij} \geqslant 0 \tag{2}$$

$$\forall_{i\in S} \sum_{j\in S} p_{ij} = 1. \tag{3}$$

Podstawowe pojęcia. III

Definicja 3

Niech ciąg zmiennych losowych $(X_n)_{n \geqslant 0}$ *będzie łańcuchem Markowa.*
Rozkład zmiennej losowej X_0 *nazywamy rozkładem początkowym.*
Łańcuch Markowa nazywamy jednorodnym (w czasie) wtedy i tylko wtedy, gdy istnieje macierz $\mathcal{P} = [p_{ij}]_{i,j \in S}$ *będąca dla każdego* n *macierzą przejścia w* n*-tym kroku.*

Jednorodne łańcuchy Markowa. I

Twierdzenie 1

Dla każdego rozkładu prawdopodobieństwa μ na S i macierzy przejścia $\mathcal{P}$ na S istnieje na pewnej przestrzeni probabilistycznej łańcuch Markowa o rozkładzie początkowym μ i macierzy przejścia $\mathcal{P}$.

Będziemy rozważać od tej pory jednorodne łańcuchy Markowa $(X_n)_{n\geqslant 0}$ o zbiorze stanów S i macierzy przejścia $\mathcal{P}$.

Jednorodne łańcuchy Markowa. II

Twierdzenie 2

Niech ciąg zmiennych losowych $(X_n)_{n\geqslant 0}$ będzie jednorodnym łańcuchem Markowa. Wówczas o ile odpowiednie prawdopodobieństwa są niezerowe, to dla dowolnych $n, m \in \mathbb{N}$, $\{s_0, \dots, s_n\} \subset S$, $\{s_0, \dots, s_n\} \subset S$ zachodzą równości

$$
\begin{aligned}
&P(\{X_k = s_k, k \in \overline{1,n}\} | \{X_0 = s_0\}) \\
&= P(\{X_{m+k} = s_k, k \in \overline{1,n}\} | \{X_0 = s_0\}), && (4)\\
&P(\{X_{n+m} = s_1\} | \{X_m = s_0\}) = P(\{X_n = s_1\} | \{X_0 = s_0\}), && (5)\\
&P(\{X_k = s_k, X_{m+l} = t_l, k \in \overline{1,m} \wedge l \in \overline{1,n}\} | \{X_0 = s_0\}) \\
&= P(\{X_k = s_k, k \in \overline{1,m}\} | \{X_0 = s_0\}) \times \\
&\times P(\{X_l = t_l, l \in \overline{1,n}\} | \{X_0 = s_m\}). && (6)
\end{aligned}
$$

Jednorodne łańcuchy Markowa. III

Definicja 4

Niech $\mathcal{P}$ będzie macierzą przejścia dla jednorodnego łańcucha Markowa. Wówczas macierz przejścia w n krokach $\mathcal{P}(n) = [p_{ij}(n)]_{i,j\in S}$ zdefiniowana jest następująco

$$\forall_{l\in\mathbb{N}} p_{ij}(n) := P(\{X_{l+n} = s_j\}|\{X_l = s_i\}), \tag{7}$$

o ile $P(\{X_l = s_i\}) > 0$.

Twierdzenie 3

Niech $\mathcal{P}$ będzie macierzą przejścia dla jednorodnego łańcucha Markowa. Wówczas $\mathcal{P}(n) = \mathcal{P}^n$.

Jednorodne łańcuchy Markowa. IV

Wniosek 1 (Równania Chapmana - Kołmogorowa)

$$\forall_{k,l \in \mathbb{N}} \mathcal{P}(k)\mathcal{P}(l) = \mathcal{P}(k+l). \tag{8}$$

Inaczej możemy to zapisać w następującej postaci

$$\forall_{k,l \in \mathbb{N}} \forall_{s_i, s_j \in S}\, p_{ij}(k+l) = \sum_m p_{im}(k) p_{mj}(l). \tag{9}$$

Klasyfikacja stanów łańcucha Markowa. I

Klasyfikacja stanów.

Definicja 5

Mówimy, że stan s_k jest osiągalny ze stanu s_j (oznaczamy $s_j \to s_k$) wtedy i tylko wtedy, gdy

$$\exists_{n\in\mathbb{N}} p_{jk}(n) > 0. \tag{10}$$

Definicja 6

Mówimy, że stany s_k i s_j się komunikują (oznaczamy $s_j \leftrightarrow s_k$) wtedy i tylko wtedy, gdy stan s_k jest osiągalny ze stanu s_j i stan s_j jest osiągalny ze stanu s_k.

Klasyfikacja stanów łańcucha Markowa. II

Definicja 7

Mówimy, że stany s_k *jest nieistotny wtedy i tylko wtedy, gdy istnieje stan* s_j *taki, że stan* s_j *jest osiągalny ze stanu* s_k *i stan* s_k *nie jest osiągalny ze stanu* s_j.

Definicja 8

Mówimy, że zbiór stanów $\mathcal{C}$ *jest zamknięty wtedy i tylko wtedy, gdy żaden stan spoza* $\mathcal{C}$ *nie da się osiągnąć wychodząc z dowolnego stanu w* $\mathcal{C}$.

Definicja 9

Pojedynczy stan s_k *tworzący zbiór zamknięty nazywamy stanem pochłaniającym.*

Klasyfikacja stanów łańcucha Markowa. III

Definicja 10

Łańcuch Markowa nazywamy nieprzywiedlnym wtedy i tylko wtedy, gdy wszystkie stany się komunikują.

Twierdzenie 4

Dla dowolnych $s_i, s_j, s_k \in S$ *zachodzi*

$$s_i \longrightarrow s_j \wedge s_j \longrightarrow s_k \Rightarrow s_i \longrightarrow s_k. \tag{11}$$

Wniosek 2

Dla dowolnych $s_i, s_j, s_k \in S$ *zachodzi*

$$s_i \longleftrightarrow s_j \wedge s_j \longleftrightarrow s_k \Rightarrow s_i \longleftrightarrow s_k. \tag{12}$$

Klasyfikacja stanów łańcucha Markowa. IV

Oznaczmy przez F_{kj} prawdopodobieństwo, że łańcuch wychodząc ze stanu s_k dotrze kiedykolwiek do stanu s_j tzn.

$$F_{kj} := P(\bigcup_{n=1}^{\infty}\{X_n = s_j\}|\{X_0 = s_k\}), \tag{13}$$

zaś przez $f_{kj}(n)$ prawdopodobieństwo, że łańcuch wychodząc ze stanu s_k dotrze do stanu s_j po raz pierwszy w n-tym kroku tzn.

$$f_{kj}(n) := P(\{X_l \neq s_j \wedge X_n = s_j \ l = 1, \ldots, n-1\}|\{X_0 = s_k\}). \tag{14}$$

wtedy

$$F_{kj} = \sum_{n=1}^{\infty} f_{kj}(n). \tag{15}$$

Klasyfikacja stanów łańcucha Markowa. V

Lemat 1

Dla dowolnej liczby naturalnej n i dowolnych stanów $s_j, s_k \in S$ zachodzi

$$p_{kj}(n) = \sum_{m=1}^{n} f_{kj}(m) p_{jj}(n-m) = \sum_{m=1}^{n} f_{kj}(n-m) p_{jj}(m), \tag{16}$$

gdzie $p_{kj}(0) = \delta_{kj}$.

Definicja 11

Niech $s_j \in S$. Stan s_j nazywamy powracającym wtedy i tylko wtedy, gdy $F_{jj} = 1$.
Stan s_j nazywamy chwilowym wtedy i tylko wtedy, gdy $F_{jj} < 1$

Klasyfikacja stanów łańcucha Markowa. VI

Definicja 12

Zmienną losową N_j (być może przyjmującą również wartość $+\infty$) zdefiniowaną równością

$$N_j = \sum_{n=1}^{\infty} \mathbb{I}_{\{X_n = s_j\}} \tag{17}$$

nazywamy zmienną liczącą ile razy proces przebywa w stanie s_j.

Klasyfikacja stanów łańcucha Markowa. VII

Twierdzenie 5

Stan s_j jest powracający wtedy i tylko wtedy, gdy

$$P(\{N_j = +\infty\}|\{X_0 = s_j\}) = 1. \tag{18}$$

Stan s_j jest chwilowy wtedy i tylko wtedy, gdy

$$P(\{N_j < +\infty\}|\{X_0 = s_j\}) = 1. \tag{19}$$

Łańcuchy chwilowy i powracający. I

Definicja 13

Średnim czasem przebywania łańcucha Markowa w stanie s_j nazywamy wielkość

$$P_j := \sum_{n=1}^{\infty} p_{jj}(n). \tag{20}$$

Następujący lemat uzasadnia tę definicję.

Lemat 2

$$P_j = \mathbb{E}(N_j | X_0 = s_j). \tag{21}$$

Łańcuchy chwilowy i powracający. II

Twierdzenie 6

Stan s_j jest powracający wtedy i tylko wtedy, gdy $P_j = +\infty$.
Stan s_j jest chwilowy wtedy i tylko wtedy, gdy $P_j < +\infty$.

Lemat 3

Jeżeli stan s_j jest stanem chwilowym, to dla każdego stanu s_i $\sum_{n=1}^{\infty} p_{ij}(n) < +\infty$, a więc $\lim_{n\to\infty} p_{ij}(n) = 0$.

Twierdzenie 7

W nieprzywiedlnym łańcuchu Markowa wszystkie stany są tego samego typu.

Łańcuchy chwilowy i powracający. III

Definicja 14

Nieprzywiedlny łańcuch Markowa nazywamy powracającym (chwilowym) wtedy i tylko wtedy, gdy jeden ze stanów, a więc i wszystkie, jest powracający (odpowiednio chwilowy).

Wniosek 3

W nieprzywiedlnym i powracającym łańcuchu Markowa mamy

$$\forall_{s_j \in S} P(\{\exists_{n \in \mathbb{N}} X_n = s_j\}) = 1, \tag{22}$$

niezależnie od rozkładu początkowego X_0.

Łańcuchy chwilowy i powracający. IV

Twierdzenie 8

Przestrzeń stanów S łańcucha Markowa można jednoznacznie przedstawić w postaci

$$S = T \cup S_1 \cup S_2 \cup \ldots, \tag{23}$$

gdzie T jest zbiorem stanów chwilowych, a S_i są nieprzywiedlnymi zamkniętymi zbiorami stanów powracających.

Stany i łańcuchy okresowe. I

Rozważmy nieprzywiedlny łańcuch Markowa i pewien jego stan $s_i \in S$. Ponieważ stan s_i komunikuje się z samym sobą, zatem istnieje liczba $n \geqslant 1$ taka, że $\mathcal{P}^n(i,i) > 0$. Niech

$$N_i := \{n \in \mathbb{N} : \mathcal{P}^n(i,i) > 0\}.$$

Wtedy

$$m, n \in N_i \Longrightarrow m + n \in N_i,$$

gdyż, dla wszystkich stanów s_i, s_j, s_k mamy następujące oszacowanie

$$\mathcal{P}^{m+n}(i,j) = \sum_{l \in S} \mathcal{P}^m(i,l)\mathcal{P}^n(l,j) \geqslant \mathcal{P}^m(i,k)\mathcal{P}^n(k,j).$$

Stąd

$$\mathcal{P}^{m+n}(i,i) \geqslant \mathcal{P}^m(i,i)\mathcal{P}^n(i,i) > 0.$$

Stany i łańcuchy okresowe. II

Definicja 15

Mówimy, że stan s_i *jest okresowy o okresie* $\nu > 1$ *wtedy i tylko wtedy, gdy* ν *jest największym wspólnym dzielnikiem liczb ze zbioru* N_i.

Twierdzenie 9

W nieprzywiedlnym łańcuchu Markowa zachodzi dokładnie jeden z następujących warunków

1. *wszystkie stany są okresowe i mają wspólny okres,*
2. *żaden ze stanów nie jest okresowy.*

Twierdzenie ergodyczne. I

Definicja 16

Niech dany będzie nieprzywiedlny łańcuch Markowa. Łańcuch ten nazywamy ergodycznym wtedy i tylko wtedy, gdy istnieje wektor π o współrzędnych $\pi_1, \dots, \pi_k$ taki, że:

1. $\pi_i > 0$ *dla wszystkich* $i \in S$,
2. *dla wszystkich* $i, j \in S$ *mamy* $\lim_{n\to\infty} \mathcal{P}^n(i,j) = \pi_j$,
3. *wektor* π *jest jedynym rozwiązaniem równania*

$$\mathcal{P}^T x = x,$$

spełniającym warunek $\sum_{i \in S} x_i = 1$.

Twierdzenie ergodyczne. II

Uwaga 2

Rozkład, o którym mowa w powyższej definicji nazywamy rozkładem stacjonarnym łańcucha ergodycznego.

Twierdzenie 10

Rozważamy nieprzywiedlny łańcuch Markowa o skończonej liczbie stanów k (tzn. $\operatorname{card} S = k$) i macierzy przejścia $\mathcal{P}$. Wówczas zachodzi dokładnie jeden z następujących warunków

1. *łańcuch jest okresowy,*
2. *łańcuch jest ergodyczny.*

Twierdzenie ergodyczne. III

Uwaga 3

Mówiąc niezbyt precyzyjnie, ergodyczność oznacza, że dla dużych n prawdopodobieństwo przejścia ze stanu i do stanu j w n krokach jest dodatnie i zależy faktycznie od stanu końcowego j, zaś nie zależy od stanu początkowego i. Prawdopodobieństwa te otrzymujemy rozwiązując odpowiedni układ równań liniowych.

Uwaga

Dowody wszystkich powyższych faktów mogą Państwo znaleźć w podręczniku J.Jakubowskiego i R. Sztencela [1]. Niektóre dowody faktów przedstawionych na tym wykładzie są dostępne na stronie [2].

Bibliografia

[1] Jakubowski Jacek i Sztencel Rafał. *Wstęp do teorii prawdopodobieństwa*. Wyd. 4. rozsz. Warszawa: Script, 2010. ISBN: 9788389716194.

[2] Ombach Jerzy i Mazur Marcin. *Rachunek prawdopodobieństwa i statystyka*. 2006. URL: http://wazniak.mimuw.edu.pl/index.php?title=Rachunek_prawdopodobie%C5%84stwa_i_statystyka (term. wiz. 28. 05. 2020).

Spis treści

Zagadnienia. I

1. Prawdopodobieństwo.
   1. Przestrzeń probabilistyczna. Własności prawdopodobieństwa.
   2. Prawdopodobieństwo warunkowe, całkowite. Wzór Bayesa.
   3. Zdarzenia niezależne. Niezależność zdarzeń, niezależność zespołowa, niezależność zdarzeń parami – ich związek.
   4. Zdarzenie zależne. Współczynnik korelacji zdarzeń i jego własności.
   5. Prawdopodobieństwo geometryczne. Paradoks Bertranda.
   6. Schemat Bernoulliego i jego uogólnienia (zagadnienia Poissona, Pascala, uogólniony schemat Bernoulliego).
2. Zmienna losowa i jej parametry.
   1. Pojęcie zmiennej losowe i rozkładu prawdopodobieństwa. Ich własności.
   2. Dystrybuanta jednowymiarowej zmiennej losowej i jej własności.
   3. Dystrybuanta wielowymiarowej zmiennej losowej i jej własności.
   4. Jednowymiarowy rozkład ciągły i dyskretny – przykłady rozkładów i ich parametry.
   5. Wielowymiarowe rozkłady ciągłe i dyskretne. Rozkłady brzegowe.

Zagadnienia. II

6. Wartość oczekiwana zmiennej losowej i jej własności.
7. Wariancja zmiennej losowej i jej własności.
8. Momenty zmiennej losowej i ich własności.
9. Inne parametry liczbowe zmiennych losowych.
10. Nierówności dla zmiennych losowych.
11. Niezależne zmienne losowe.
12. Kowariancja zmiennych losowych. Macierz kowariancji i korelacji.
13. Zmienne losowe nieskorelowane. Niezależne zmienne losowe, a nieskorelowane.

3. Zbieżność ciągów zmiennych losowych.
   1. Zbieżność prawie na pewno i jej własności.
   2. Zbieżność prawie na pewno, a według prawdopodobieństwa.
   3. Zbieżność według prawdopodobieństwa i jej własności.
   4. Zależności między różnymi rodzajami zbieżności zmiennych losowych.
4. Prawa wielkich liczb i centralne twierdzenie graniczne.
   1. Pojęcie słabego praw wielkich liczb. Warunki dostateczne na jego zachodzenie.

Zagadnienia. III

2. Pojęcie mocnego praw wielkich liczb. Twierdzenia Kołmogorowa.
3. Twierdzenia Moivre'a-Laplace'a.
4. Schemat serii, znormalizowany schemat serii.
5. Warunek Lindeberga i jego własności.
6. Twierdzenie Lindeberga - Lévy'ego.

5. Łańcuchy Markowa.
   1. Podstawowe pojęcie dotyczące łańcuchów Markowa.
   2. Klasyfikacja stanów.
   3. Klasyfikacja łańcuchów.
   4. Twierdzenie ergodyczne.
6. Podstawowe pojęcia statystyczne.
7. Elementy statystyki opisowej.
   1. Podstawowe pojęcia statystyki opisowej.
   2. Miary przeciętne.
   3. Miary zmienności.

Zagadnienia. IV

4. Miary asymetrii.
5. Miary koncentracji.

8. Próba losowa i statystyka z próby.
9. Rozkłady statystyk.
   1. Rozkład średniej i różnicy średnich.
   2. Rozkład wariancji i ilorazu wariancji.
   3. Rozkłady graniczne niektórych statystyk.
10. Teoria estymacji.
    1. Podstawy teorii estymacji – podstawowe pojęcia i ich rodzaje
    2. Rodzaje estymatorów i ich własności.
    3. Nierówność Rao - Cramera.
11. Estymacja punktowa.
    1. Metoda momentów konstrukcji estymatorów.
    2. Metoda największej wiarogodności konstrukcji estymatorów.

Zagadnienia. V

3. Metoda najmniejszych kwadratów konstrukcji estymatorów.

12. Estymacja przedziałowa.
    1. Podstawowe pojęcia związane z estymacją przedziałową.
    2. Przedział ufności dla średniej m w populacji normalnej i nieznanym rozkładzie
    3. Przedział ufności dla wariacji σ^2 dla populacji normalnej o nieznanych wartości oczekiwanej i odchylenia standardowego.
    4. Przedział ufności dla parametru frakcji.
    5. Estymacja przedziałowa – problem minimalizacji próby.
13. Testowanie hipotez.
    1. Podstawowe pojęcia.
    2. Testy statystyczne (etapy konstrukcji, błędy, itd.).
14. Parametryczne testy istotności.
    1. Test istotności dla wartości średniej populacji generalnej.
    2. Test istotności dla wariancji.

Zagadnienia. VI

15. Nieparametryczne testy istotności.
    1. Klasyfikacja testów.
    2. Przykładowe test zgodności.
    3. Test niezależności chi-kwadrat.
16. Korelacja cech.
    1. Miary zależności nieliniowej oparte na statystyce chi-kwadrat.
    2. Współczynnik korelacji liniowej Pearsona.
    3. Współczynnik korelacji rang Spearmana.
    4. Stosunki korelacyjne.
17. Klasyczny model regresji liniowej.
    1. Sformułowanie klasycznego modelu regresji liniowej i klasycznego modelu normalnej regresji liniowej.
    2. Estymacja parametrów funkcji regresji.
    3. Dokładność dopasowania prostej metodą najmniejszych kwadratów.
    4. Twierdzenie Gaussa-Markowa.

Zagadnienia. VII

5. Wnioskowanie o klasycznym modelu normalnej regresji liniowej.
6. Macierzowe ujęcie modelu regresji liniowej.
7. Klasyczny model regresji liniowej z wieloma zmiennymi niezależnymi.
8. Inne niż liniowe modele regresji.
9. Założenia modelu regresji liniowej i ich testowanie.

18. Jednoczynnikowa analiza wariancji.
    1. Podstawowe pojęcia.
    2. Założenia analizy wariancji.
    3. Testy post-hoc.