

Wrażliwość rozkładu stacjonarnego łańcucha Markowa na zaburzenia

Barbara Rolska

19.03.2025

Spis treści

1 Podstawowe pojęcia - łańcuchy Markowa

2 Współczynniki uwarunkowania

Łańcuch Markowa

- Łańcuch Markowa to proces określony na przestrzeni stanów $\mathcal{X}$ - w pracy rozważamy tylko przeliczalne przestrzenie stanów.
- Jeśli w pierwszym kroku znajdujemy się w stanie $x \in \mathcal{X}$, to w następna pozycja wybierana jest zgodnie z rozkładem prawdopodobieństwa $P(x, \cdot)$.
- Macierz P (macierz przejścia) określamy następująco: dla $x, y \in \mathcal{X}$ element $P(x, y)$ oznacza prawdopodobieństwo przejścia ze stanu x do stanu y w kolejnym kroku.
- Dla ustalonego $x \in \mathcal{X}$ mamy:

$$\sum_{y \in \mathcal{X}} P(x, y) = 1.$$

Łańcuch Markowa - definicja

Definicja

Ciąg zmiennych losowych $\{X_n\}_{n \in \mathbb{N}}$ o wartościach w przestrzeni stanów $\mathcal{X}$ nazywamy **łańcuchem Markowa**, wtedy i tylko wtedy, gdy dla każdego $n \in \mathbb{N}$ i każdego ciągu $x_0, x_1, \dots, x_n \in \mathcal{X}$ mamy

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0) = \mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}),$$

jeśli $\mathbb{P}(X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0) > 0$. Własność tę nazywamy **własnością Markowa**.

Definicja

Rozkładem początkowym μ nazywamy rozkład zmiennej losowej X_0 .

Łańcuch Markowa - definicja

Uwaga

Oznaczmy rozkład zmiennej losowej X_t jako μ_t , czyli $\mu_t(x) = \mathbb{P}(X_t = x)$. Wtedy dla każdego $y \in \mathcal{X}$

$$\mu_{t+1}(y) = \sum_{x \in \mathcal{X}} \mathbb{P}(X_t = x)P(x, y) = \sum_{x \in \mathcal{X}} \mu_t(x)P(x, y).$$

Inaczej

$$\mu_{t+1} = \mu_t P,$$

czyli

$$\mu_t = \mu_0 P^t.$$

Rozkład stacjonarny

Rozważamy granicę $\pi = \lim_{t \rightarrow \infty} \mu_t$.

Definicja

Mówimy, że rozkład prawdopodobieństwa $\{\pi_j\}_{j \in \mathcal{X}}$ jest **rozkładem stacjonarnym** dla łańcucha Markowa o macierzy przejścia P , gdy

$$\pi = \pi P.$$

Nieredukowalność

Definicja

Stan y nazywamy **osiągalnym** ze stanu x , jeśli zachodzi $P^n(x, y) > 0$ dla pewnego n , czyli da się przejść z x do y . Tę relację oznaczamy $x \rightarrow y$. Stany x i y nazywamy **wzajemnie komunikującymi się**, jeżeli $x \rightarrow y$ i $y \rightarrow x$.

Niereducowalność

Definicja

Stan y nazywamy **osiągalnym** ze stanu x , jeśli zachodzi $P^n(x, y) > 0$ dla pewnego n , czyli da się przejść z x do y . Tę relację oznaczamy $x \rightarrow y$. Stany x i y nazywamy **wzajemnie komunikującymi się**, jeżeli $x \rightarrow y$ i $y \rightarrow x$.

Definicja

Łańcuch Markowa nazywamy **niereducowalnym**, jeśli każde dwa stany z przestrzeni stanów wzajemnie się komunikują, to znaczy z każdego stanu można przejść do dowolnego innego (niekoniecznie w jednym kroku).

Niereducowalność

Definicja

Stan y nazywamy **osiągalnym** ze stanu x , jeśli zachodzi $P^n(x, y) > 0$ dla pewnego n , czyli da się przejść z x do y . Tę relację oznaczamy $x \rightarrow y$. Stany x i y nazywamy **wzajemnie komunikującymi się**, jeżeli $x \rightarrow y$ i $y \rightarrow x$.

Definicja

Łańcuch Markowa nazywamy **niereducowalnym**, jeśli każde dwa stany z przestrzeni stanów wzajemnie się komunikują, to znaczy z każdego stanu można przejść do dowolnego innego (niekoniecznie w jednym kroku).

Twierdzenie

Każdy niereducowalny łańcuch Markowa posiada jedyny rozkład stacjonarny.

Spis treści

1 Podstawowe pojęcia - łańcuchy Markowa

2 Współczynniki uwarunkowania

Zarys problemu

- Niech $\{X_n\} \equiv \{X_n, n = 0, 1, 2, \dots\}$ będzie ergodycznym łańcuchem Markowa z dyskretną przestrzenią stanów $J = \{0, 1, 2, \dots\}$, macierzą przejścia $P = (p_{i,j}, i, j \in J)$ i rozkładem stacjonarnym $\pi = (\pi_i, i \in J)$.
- $\tilde{P} = (\tilde{p}_{i,j}, i, j \in J)$ - **zaburzenie macierzy P** (stochastyczna, ergodyczna)
 $\tilde{\pi} = (\tilde{\pi}_i, i \in J)$ - rozkład stacjonarny $\tilde{P}$
 $E = P - \tilde{P}$ - **macierz wielkości zaburzeń.**
 Inaczej: $\tilde{P} = P - E$

Przykład

$$\begin{aligned}
 \tilde{P} &= \begin{pmatrix} 0,74 & 0,11 & 0 & 0 & 0 & 0 & 0 & 0,15 \\ 0 & 0,689 & 0 & 0 & 0,011 & 0 & 0 & 0,30 \\ 0 & 0 & 0 & 0,40 & 0 & 0 & 0 & 0,60 \\ 0 & 0 & 0 & 0,669 & 0,011 & 0 & 0 & 0,32 \\ 0 & 0 & 0 & 0 & 0,912 & 0 & 0 & 0,088 \\ 0 & 0 & 0 & 0 & 0 & 0,74 & 0 & 0,26 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,87 & 0,13 \\ 0,15 & 0 & 0,047 & 0 & 0 & 0,055 & 0,27 & 0,478 \end{pmatrix} - \begin{pmatrix} 0,01 & 0,01 & 0 & 0 & 0 & 0 & -0,005 & -0,015 \\ 0 & 0,005 & 0 & 0 & 0 & -0,01 & 0 & 0,005 \\ 0 & 0 & -0,02 & 0,02 & 0 & 0 & -0,015 & 0,015 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0,07 & 0 & 0 & 0 & 0 & 0 & 0 & 0,07 \\ 0,03 & 0 & 0 & 0 & 0 & 0,03 & -0,07 & 0,01 \end{pmatrix} \\
 &= \begin{pmatrix} 0,73 & 0,1 & 0 & 0 & 0 & 0 & 0,005 & 0,165 \\ 0 & 0,684 & 0 & 0 & 0,011 & 0,01 & 0 & 0,295 \\ 0 & 0 & 0,02 & 0,38 & 0 & 0 & 0,015 & 0,585 \\ 0 & 0 & 0 & 0,669 & 0,011 & 0 & 0 & 0,32 \\ 0 & 0 & 0 & 0 & 0,912 & 0 & 0 & 0,088 \\ 0 & 0 & 0 & 0 & 0 & 0,74 & 0 & 0,26 \\ 0,07 & 0 & 0 & 0 & 0 & 0 & 0,87 & 0,06 \\ 0,12 & 0 & 0,047 & 0 & 0 & 0,025 & 0,34 & 0,468 \end{pmatrix}
 \end{aligned}$$

Problem

Szukamy ograniczenia (zależnego od P) dla $\|\pi - \tilde{\pi}\|_p$ postaci:

$$\|\pi - \tilde{\pi}\|_p \leq \kappa \|E\|_q$$

κ nazywamy **współczynnikiem uwarunkowania** (condition number).

Normy dla macierzy i wektorów

$$\|\pi - \tilde{\pi}\|_1 = \sum_{j \in J} |\pi_j - \tilde{\pi}_j|$$

$$\|\pi - \tilde{\pi}\|_\infty = \sup_{j \in J} |\pi_j - \tilde{\pi}_j|$$

$$\|B\|_\infty = \sup_{i \in J} \sum_{j \in J} |b_{i,j}|$$

$$\|B\|_{\text{sup}} = \sup_{i,j \in J} |b_{i,j}|.$$

Macierz fundamentalna

Definicja

Macierz fundamentalna Z dla łańcucha Markowa o macierzy przejścia P oraz z rozkładem stacjonarnym π jest zdefiniowana następująco:

$$Z = (I - P + e\pi)^{-1},$$

gdzie e to wektor jednostkowy $(1, 1, 1, 1, \dots, 1)^T$.

Oznaczenie

Wprowadzamy oznaczenie:

$$A = I - P$$

Wtedy $A^\#$ to grupowa odwrotność macierzy A , czyli taka, która spełnia:

$$AA^\#A = A, \quad A^\#AA^\# = A^\#, \quad AA^\# = A^\#A.$$

Współczynnik uwarunkowania 1

$$\pi - \tilde{\pi} = \tilde{\pi}EZ$$

$$\|\pi - \tilde{\pi}\|_1 \leq \|Z\|_\infty \|E\|_\infty$$

$$\kappa_1 = \|Z\|_\infty - \text{Schweitzer 1968}$$

Współczynnik uwarunkowania 2

$$\pi - \tilde{\pi} = \tilde{\pi}EA^\#$$

$$\|\pi - \tilde{\pi}\|_1 \leq \|A^\#\|_\infty \|E\|_\infty$$

$$\kappa_2 = \|A^\#\|_\infty - \text{Meyer 1980}$$

Współczynniki uwarunkowania 3 i 4

1

$$\|\pi - \tilde{\pi}\|_{\infty} \leq \kappa_3 \|E\|_{\infty}$$

$$\kappa_3 = \frac{\max_j \left(a_{jj}^{\#} - \min_i a_{ij}^{\#} \right)}{2} \text{ - Haviv 1984, Kirkland 1998}$$

2

$$\|\pi - \tilde{\pi}\|_{\infty} \leq \kappa_4 \|E\|_{\infty}$$

$$\kappa_4 = \max_{i,j} \left| a_{ij}^{\#} \right| \text{ - Funderlic, Meyer 1986}$$

Współczynniki uwarunkowania 5 i 6

Współczynnik ergodyczności:

$$\tau_1(P) = \frac{1}{2} \max_{i,j} \sum_{k=1}^n |p_{ik} - p_{jk}|$$

$$\tau_1(B) = \sup_{ye^T=0} \frac{\|yB\|_1}{\|y\|_1}$$

1

$$\|\pi - \tilde{\pi}\|_{\infty} \leq \kappa_5 \|E\|_{\infty}$$

$$\kappa_5 = \frac{1}{1 - \tau_1(P)} - \text{Seneta 1988}$$

2

$$\|\pi - \tilde{\pi}\|_1 \leq \kappa_6 \|E\|_{\infty}$$

$$\kappa_6 = \tau_1(A^{\#}) = \tau_1(Z) - \text{Seneta 1991}$$

Współczynnik uwarunkowania 7

$$\|\pi - \tilde{\pi}\|_{\infty} \leq \kappa_7 \|E\|_{\infty}$$

$$\kappa_7 = \frac{1}{2} \min_j \|A_{(j)}^{-1}\|_{\infty} - \text{Ipsen, Meyer 1994}$$

$A_{(j)}$ jest podmacierzą A uzyskaną przez usunięcie j -tego wiersza i kolumny z A

Współczynnik uwarunkowania 8

Średni czas przejścia ze stanu i do j oznaczamy jako m_{ij} .

$$\|\pi - \tilde{\pi}\|_{\infty} \leq \kappa_8 \|E\|_{\infty}$$

$$\kappa_8 = \frac{1}{2} \max_j \left(\frac{\max_{i \neq j} m_{i,j}}{m_{j,j}} \right) - \text{Cho, Meyer 2000}$$