Poznámky k Markovovým řetězcům

(Diskrétní) Markovův řetězec je posloupnost náhodných veličin $X_0, X_1, \dots$ se stejným (konečným) oborem hodnot obvykle ve tvaru $\{1, \dots, k\}$. Hodnoty se nazývají stavy.

Řetězec splňuje tuto podmínku

$$P\Big(X_{n+1} = j \Big| X_n = i,\ X_{n-1} = i_{n-1} \dots,\ X_0 = i_0\Big) = P\Big(X_{n+1} = j \Big| X_n = i\Big) = p_{i,j}$$

pro všechna $n = 0, 1, \dots$ a všechny stavy $i, j, i_0, \dots, i_{n-1} \in \{1, \dots, k\}$.

Hodnoty $p_{i,j}$ jsou prvky tzv. *matice přechodu* $\mathbf{P}$ typu $k \times k$, která NEZÁVISÍ na n.

Podmínka říká, že podmíněná pravděpodobnost budoucího stavu j (v čase $n+1$) je (při zadání nějaké posloupnosti stavů od času 0 až do času n) plně určena tím posledním z nich, tj. stavem i v čase n.

Odsud můžeme hned odvodit následující:

$$P\Big(X_0 = i_0\ , \dots,\ X_n = i_n\Big) =$$

$$= P\Big(X_n = i_n \Big| X_0 = i_0\ , \dots,\ X_{n-1} = i_{n-1}\Big) \cdot P\Big(X_0 = i_0\ , \dots,\ X_{n-1} = i_{n-1}\Big) =$$

$$= P\Big(X_n = i_n \Big| X_{n-1} = i_{n-1}\Big) \cdot P\Big(X_0 = i_0\ , \dots,\ X_{n-1} = i_{n-1}\Big) =$$

$$= p_{i_{n-1}, i_n} \cdot P\Big(X_0 = i_0\ , \dots,\ X_{n-1} = i_{n-1}\Big)$$

Iterací dostaneme postupně, že

$$P\Big(X_0 = i_0\ , \dots,\ X_n = i_n\Big) = P\Big(X_0 = i_0\Big) \cdot p_{i_0,i_1} \cdot p_{i_1,i_2} \cdots p_{i_{n-1},i_n}\ ,$$

kde $p_{j,\ell}$ je prvek matice $\mathbf{P}$ v j-tém řádku a ℓ-tém sloupci.

Co se dá ještě podobně odvodit:

- pro každé $m = 0, 1, \dots$ je
  $$X_m, X_{m+1}, \dots$$
  opět Markovův řetězec (což vypadá i dost intuitivně)

- $$P(X_{t+n} = i_n,\ X_{t+n-1} = i_{n-1},\ \dots,\ X_{t+1} = i_1 | X_t = i_0) = p_{i_0,i_1} \dots p_{i_{n-1},i_n}$$
  pro všechny patřičné proměnné. Speciálně to znamená, že tato hodnota vůbec nezávisí na t. Tuhle podmíněnou pravděpodobnost proto označíme jako
  $$P(i_0 \to i_1 \to \dots \to i_n)$$
  protože vyjadřuje pravděpodobnost, toho, že (za předpokladu že jsme nyní v i_0) se dostaneme ze stavu i_0 po n krocích do stavu i_n tak, že postupně projdeme stavy i_1 až i_{n-1}.

- podobně i tato podmíněná pravděpodobnost nezávisí na t:
  $$P(X_{t+n} = j | X_t = i) = \sum_{i_1,\dots,i_{n-1}} p_{i,i_1} \dots p_{i_{n-1},j} = (\mathbf{P}^n)_{i,j}$$
  a je, jak vidíme, rovna prvku na pozici (i, j) v n-té mocnině matice $\mathbf{P}$. Tuto podmíněnou pravděpodobnost označíme analogicky $P\Big(\underbrace{i \to \cdot \to \dots \to \cdot \to j}_{n \text{ kroků}}\Big)$ tj. tím vlastně procházíme všechny cestu z i do j délky n.

Definice: Komponenta K Markovova řetězce je komponenta silné souvislosti příslušného orientovaného grafu, ve kterém jsme ponechali pouze trvalé stavy.

Tj. K je taková neprázdná podmnožina množiny trvalých stavů, že

1. libovolné dva různé stavy $i, j \in K$ jsou navzájem propojené orientovanými cestami,
2. jestliže stav $i \in K$ je propojený orientovanou cestou se nějakým trvalým stavem k, pak tento stav k musí také ležet v komponentě K.

Definice: Nechť K je komponenta a množiny $M_1, \ldots, M_k$ pro $k \geq 2$ tvoří její rozklad (tj. jsou navzájem disjunktní a jejich sjednocením je celá komponenta K). Řekneme, že tento rozklad je *cyklický* (a délky d), pokud

- Uvnitř těchto množin M_i nevedou mezi stavy šipky (pro všechna $i = 1, \ldots, d$).
- Naopak všechny šipky z dané množiny M_i musí směrovat do množiny M_{i+1} pro $i = 1, \ldots, d-1$ a šipky z M_d směřují zase zpět do M_1.

Definice: Řekneme, že komponenta K má periodu π, jestliže $\pi \in \mathbb{N}$ je největší takové číslo, že existuje cyklický rozklad komponenty K délky π (v tom případě je nutně $\pi \geq 2$). Jestliže žádný takový rozklad neexistuje, definujeme $\pi := 1$ (případně se říká, že komponenta je pak aperiodická).

Perioda $\pi(i)$ (trvalého) stavu i je definována jako

$$\pi(i) := \gcd(\text{"délek všech uzavřených cest obsahujících stav } i\text{"})$$

Věta (o periodě komponenty): Nechť K je komponenta s alespoň dvěma stavy a π její perioda. Pak

(i) $\pi = \gcd(\text{"délek všech uzavřených cest v komponentě } K\text{"})$

(ii) Pro každý stav $i \in K$ je $\pi(i) = \pi$ (speciálně tedy všechny stavy mají stejnou periodu).

(iii) Jestliže $N_1, \ldots, N_d$ je cyklický rozklad komponenty K, pak d dělí π. Dále si pro toto d definujme relaci $\sim_d$ pro stavy $a, b \in K$ jako

$a \sim_d b \Leftrightarrow$ stav a je propojený se stavem b orientovanou cestou (buď tam nebo zpět), jejíž délka je dělitelná d.

Pak $\sim_d$ je relace ekvivalence na K a její rozkladové třídy jsou právě množiny $N_1, \ldots, N_d$.

Postup jak najít hodnotu periody π dané komponenty:

- vezmeme (konečně mnoho) nějakých uzavřených cest v dané komponentě a spočítáme

  $$k = \gcd(\text{"délky těchto konečně mnoha uzavřených cest"})\ .$$

  Pak π musí dělit k.

- Jestliže $k = 1$, pak $\pi = 1$. Jestliže $k \geq 2$, pak projdeme všechny jeho dělitele d (kterých je konečně mnoho) a pro každý takový dělitel d zkusíme zjistit, jestli relace $\sim_d$ je ekvivalence (na K) a jestli rozkladové třídy této ekvivalence tvoří cyklický rozklad komponenty K.

  Rozkladové třídy zde vytváříme tak, že k sobě budeme sdružovat stavy, které jsou spojeny cestami, jejichž délky jsou dělitelné d.

  No a největší číslo d s touto vlastností bude hledaná perioda π.

Věta (o odvozeném řetězci): Nechť Markovův řetězec má jedinou komponentu K s periodou $\pi \geq 2$ a maticí přechodu $\mathbf{P}$. Nechť množiny $M_1, \ldots, M_\pi$ tvoří cyklický rozklad této komponenty.

Nový Markovovův řetězec daný aplikováním π kroků je pak určený maticí $\mathbf{P}^\pi$ a jeho komponenty budou právě množiny $M_1, \ldots, M_\pi$, z nichž každá bude mít periodu 1.

Věta (o stacionárních rozděleních): Pro Markovův řetězec s komponentami $K_1, \ldots, K_\ell$ a jejich periodami $\pi_1, \ldots, \pi_\ell$ platí, že:

(i) Každá komponenta K_i má *jediné* své stacionární rozdělení $\boldsymbol{p}_i$ takové, že jeho hodnoty jsou nenulové na stavech z K_i a nulové na stavech mimo K_i. Libovolné další stacionární rozdělení $\boldsymbol{p}$ na celém řetězci je pak nějakou konvexní kombinací těchto speciálních rozdělení, tedy existují $\alpha_1, \ldots, \alpha_\ell \geq 0$ taková, že $\sum_{i=1}^{\ell} \alpha_i = 1$ a $\boldsymbol{p} = \sum_{i=1}^{\ell} \alpha_i \boldsymbol{p}_i$ (a naopak libovolná konvexní kombinace stacionárních rozdělení je opět stacionární).

(ii) Každé počáteční rozdělení konverguje k *nějakému* stacionárnímu $\Leftrightarrow$ každá komponenta K_i tohoto řetězce má periodu $\pi_i = 1$.

(iii) Každé počáteční rozdělení konverguje k *jedinému* stacionárnímu $\Leftrightarrow$ řetězec má jedinou komponentu s periodou 1 (tedy je nerozložitelný a aperiodický, neboli *ergodický*).

Jak hledat asymptotická rozdělení Markovova řetězce M, kde každá komponenta je neperiodická (tj. případ (ii) z Věty o stacionárních rozděleních):

- Množinu všech trvalých stavů řetězce M rozdělíme na komponenty $K_1, \ldots, K_\ell$.
- Vytvoříme nový Markovovův řetězec M', který vznikne z původního takto
  - přechodné stavy řetězce M' budou stejné jako přechodné stavy původního M,
  - trvalé stavy řetězce M' vzniknou nahrazením každé komponenty K_i jediným absorbčním stavem a_i.
  - šipky mezi přechodnými stavy v M' budou stejné jako v M
  - šipky z přechodných stavů do absorbčních stavů budou v M' tyto: Nechť j je teď přechodný stav (původního řetězce M) a $q_1, \ldots, q_m$ jsou hodnoty všech šipek, které vedou z j do nějakého trvalého stavu v K_i. V novém Markovově řetězci M' nyní všechny tyto šipky nahradíme šipkou jedinou, která bude vést z j do a_i a přiřadíme jí hodnotu $q = \sum_{k=1}^{m} q_k$.
- Mějme nyní počáteční rozdělení $\mathbf{p}(0)$ na řetězci M. Z něj vytvoříme počáteční rozdělení $\mathbf{p}'(0)$ na řetězci M' tak, že hodnoty pravděpodobností v přechodných stavech ponecháme a hodnoty pravděpodobností v rámci dané komponenty K_i sečteme (a výsledek přiřadíme stavu a_i).
- Nyní pro nový Markovův řetězec M' určíme asymptotické rozdělení pravděpodobnosti $\mathbf{p}'(\infty)$ známou metodou pro řetězec, kde jsou JEN přechodné a absorbční stavy.
- Nechť α_i je pravděpodobnost (tohoto asymptotického rozdělení $\mathbf{p}'(\infty)$) pro absorbční stav a_i v M' (který odpovídá komponentě K_i původního řetězce M). Najdeme (jediné) stacionární řešení $\mathbf{p}_i$ pro komponentu K_i původního řetězce takové, že $\mathbf{p}_i$ je nenulové na dané komponentě K_i a nulové všude jinde na původním řetězci M.

  Pak asymptotickým rozdělením $\mathbf{p}(\infty)$ původního řetězce M pro počáteční rozdělení $\mathbf{p}(0)$ je tato konvexní kombinace

$$\mathbf{p}(\infty) = \sum_{i=1}^{\ell} \alpha_i \mathbf{p}_i \ .$$