Marsovske verige

Živa Flego, Kaja Rajter in Lesley Zore
Mentor: Žan Hafner Petrovski

Povzetek

V projektu smo si zastavili MaRSovski problem, ki smo ga rešili z uporabo priročne interpretacije stacionarne porazdelitve Markovske verige. Sproti smo obnovili znanje o verjetnosti in spoznali pojem matrike. Seznanjeni z osnovnimi pojmi smo definirali Markovske verige ter govorili o izreku, ki velja za krepko povezane neperiodične Markovske verige. Izrek nam pomaga razumeti limitno obnašanje verige, to pa nam je dalo ključen vpogled v rešitev zastavljenega problema.

1 Uvod

Oglejmo si problem [2], ki ga bomo kasneje rešili s prevedbo na Markovske verige.

Problem 1. *Marsovce in Zemljane opazujemo, kako se en za drugim pomikajo v vrsti. Verjetnost, da za Zemljanom stoji Marsovec, je* 0.75 *in* 0.25, *da za njim stoji Zemljan. Verjetnost, da za Marsovcem stoji Marsovec, je* 0.8 *in* 0.2, *da za njim stoji Zemljan. Kolikšen delež vseh pohodnikov na stezi je Marsovcev in kolikšen Zemljanov?*

Brez škode za splošnost si lahko mislimo, da je Marsovcev in Zemljanov neskončno in da smo trenutno opazili Marsovca. Ob opazovanju naslednjega osebka bomo z verjetnostjo 0.8 opazili Marsovca in z verjetnostjo 0.2 Zemljana. Zakaj tako razmišljanje ne škodi splošnosti, bomo videli v izreku.

2 Verjetnost

Za verjetnost poznamo več definicij, ena izmed njih se imenuje tudi klasična definicija, ki je navedena spodaj [3].

Definicija 1. *Če ima poskus končno število enako verjetnih izidov, je* ***verjetnost*** *količnik med številom ugodnih izidov oz. izbranih dogodkov in številom vseh možnih izidov. Naj množica* Ω *predstavlja vse možne izide in naj bo* $A \subseteq \Omega$ *dogodek. Verjetnost, da se zgodi dogodek* A*, zapišemo kot*

$$\textit{verjetnost, da se zgodi } A = \frac{\textit{št. ugodnih izidov}}{\textit{št. vseh izidov}}.$$

V kolikor označimo množico ugodnih izidov oziroma dogodkov z A, verjetnost dogodka A s $P(A)$ in množico prostora vseh mogočih dogodkov z Ω, dobimo naslednji zapis zgornje enačbe

$$P(A) = \frac{|A|}{|\Omega|}.$$

Definicija nam pove tudi, da je *dogodek* pojav, ki se pri poskusu (npr. metu kocke) lahko zgodi. Dogodek je namreč podmnožica množice vseh možnih izidov. Dogodek je lahko gotov, nemogoč ali slučajen. Verjetnost gotovega dogodka je 1, nemogočega dogodka 0, verjetnost slučajnega dogodka pa se giblje med številoma 0 in 1.

Najbolj klasičen primer, ki ga vzamemo za prikaz oziroma izračun verjetnosti, je *met kocke*.

Zgled 1. *Vržemo kocko. Kolikšna je verjetnost, da vržemo število 4? Množico vseh izidov predstavlja* $\Omega = \{1, 2, 3, 4, 5, 6\}$*, dogodek* A *je met števila* 4*. Vidimo, da je verjetnost, da bomo vrgli* 4*:*

$$P(A) = \frac{1}{6}.$$

Ob tem opazimo, da je dogodek, da vržemo eno izmed števil od 1 *do* 6*, gotov in dogodek, da bomo vrgli hkrati* 4 *in da bo število liho, nemogoč.*

Kadar govorimo o verjetnosti, moramo nujno omeniti tudi relacijo med dogodki. Dogodki so med seboj lahko odvisni ali neodvisni.

Definicija 2. *Dogodka* A *in* B *sta* ***neodvisna****, če velja*

$$P(A \cap B) = P(A)P(B).$$

Oglejmo si, kaj to pomeni pri metu kocke.

Zgled 2. *Kolikšna je verjetnost, da v dveh metih kocke vržemo število* 3 *v prvem metu, nato pa sodo število? V tem primeru je:*

- *dogodek* A *v* 1*. metu vržemo število* 3*,*
- *dogodek* B *v* 2*. metu vržemo sodo število.*

Očitno velja

$$P(A) = \frac{1}{6}, \quad P(B) = \frac{3}{6}.$$

Ker sta prvi in drugi met neodvisna, velja tudi

$$P(A \cap B) = \frac{1}{6} \cdot \frac{3}{6} = \frac{3}{36} = \frac{1}{12}.$$

Oglejmo si še en zgled.

Zgled 3. *Ali sta sledeča dogodka* A *in* B *neodvisna?*

- A *dogodek, da je vsota dveh metov* 6*,*
- B *dogodek, da je prvi met* 4*.*

Velja

$$P(A) = \frac{5}{36}, \quad P(B) = \frac{1}{6}$$

in če preštejemo vse možne ugodne izide, je

$$P(A \cap B) = \frac{1}{36}.$$

Dobimo torej

$$P(A \cap B) \neq P(A)P(B),$$

zato sta dogodka odvisna.

Vzemimo zdaj namesto A dogodek A', da je vsota dveh metov 7. Dobimo

$$P(A') = \frac{6}{36}, \quad P(A' \cap B) = \frac{1}{36}.$$

Dogodka sta začuda neodvisna. V tem primeru namreč ni nujno, da je prvi met manj kot 6 kot pri dogodku A. Pri A' lahko vržemo karkoli pri prvem metu in nam to ne pove nič o končni vsoti, pri prejšnjem primeru pa nam.

Definicija 3. ***Pogojno verjetnost****, da se zgodi A, če se zgodi B, izračunamo po predpisu*

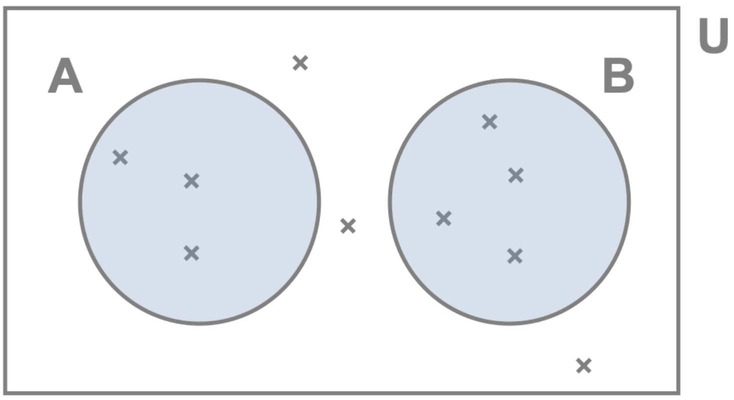
$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Omenimo še pojem disjunktnosti dogodkov.

Definicija 4. *Dogodka A in B sta* ***disjunktna****, če je njun presek prazna množica*

$$A \cap B = \emptyset.$$

Vennov diagram za dva neodvisna dogodka vidimo na sliki 1.



Slika 1: Shematični prikaz disjunktnih dogodkov A in B [4].

Kot preprost zgled navedimo met kocke, kjer je dogodek A met števila 5 in dogodek B met števila 4. Verjetnost, da se oba dogodka zgodita istočasno, je enaka 0.

3 Nekaj o matrikah

Matrika je pravokotna razpredelnica elementov oziroma števil. Uporabna je za zapis podatkov, a ko na matrikah definiramo še neke operacije, npr. seštevanje in množenje, opazimo širino njihove namembnosti.

Matrike so torej tabele sestavljene iz m vrstic in n stolpcev. V našem primeru bodo elementi matrike z intervala $[0, 1]$, zato lahko za matriko A velikosti $m \times n$ pišemo $A \in [0, 1]^{m \times n}$.

Z matrikami lahko izvajamo različne matematične operacije. Kot prvo pokažimo, kako seštejemo matriki $A, B \in \mathbb{R}^{m \times n}$. Naj bosta

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix},$$

potem je

$$A + B = \begin{bmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{bmatrix}.$$

Matrike torej seštevamo tako, da seštejemo istoležne elemente.

Nekoliko zahtevnejša operacija pa je množenje. Matrike množimo tako, da med seboj skalarno množimo ustrezne vrstice in stolpce matrik. Naj bosta matriki A in B taki kot prej. Potem je produkt AB oblike

$$C = AB = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mn} \end{bmatrix}$$

kjer vrednosti c_{ij} izračunamo po enačbi:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Vrednost c_{ij} je torej skalarni produkt i-te vrstice in j-tega stolpca.

4 Markovske verige

Ruski matematik Andrej Andrejevič Markov (1856-1922) je deloval na področjih teorije števil, verjetnosti in matematične analize. Po njem so poimenovani številni matematični pojmi, med drugim tudi Markovske verige.

Definicija 5. *Zaporedje slučajnih spremenljivk* $X_0, X_1, X_2, \ldots$ *je* ***Markovska veriga****, če upošteva lastnost Markova. To pomeni, da mora za poljuben* $n \in \mathbb{N}$ *veljati*

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = P(X_n = x_n | X_{n-1} = x_{n-1}).$$

Za vsak $n \in \mathbb{N}$ definicija pravi, da je verjetnost, da bomo po n korakih v stanju x_n, odvisna le od stanja x_{n-1}, vse predhodne korake pa lahko pozabimo. Zgodovina prehodov med stanji pred prihodom v trenutno stanje torej ne sme vplivati na verjetnost prehoda v naslednje stanje.

V definiciji smo omenili tudi **slučajne spremenljivke**. To so matematični objekti, katerih vrednost je odvisna od nekega slučajnega procesa. Ta proces bomo v primeru Markovskih verig hitro razumeli.

Markovsko verigo predstavimo z množico možnih stanj S, prehodno matriko P in začetno porazdelitvijo μ_0. V tem primeru bomo imeli zaporedje slučajnih spremenljivk $X_0, X_1, X_2, \ldots$, ki bodo odvisne od slučajnega procesa prehajanja po stanjih iz S. Za nek $i \in \mathbb{N}$ slučajna spremenljivka X_i torej pove, v katerem stanju iz S se nahajamo po i korakih, verjetnost, da pridemo v neko stanje, pa je odvisna le od predhodnega stanja.

Povejmo še, kaj je verjetnostna porazdelitev neke slučajne spremenljivke.

Definicija 6. ***Verjetnostna porazdelitev*** *določa verjetnost, da slučajna spremenljivka zavzame neko vrednost. Za slučajno spremenljivko* X *z vrednostmi v* $\{1, 2, \ldots, n\}$ *in porazdelitvijo* $\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_n)$*,* $\mu_k \in [0, 1]$ *za* $k \in \{1, 2, \ldots, n\}$ *in* $\sum_{k=1}^{n} \mu_k = 1$ *velja*

$$P(X = k) = \mu_k, \quad k \in \{1, 2, \ldots, n\}. \tag{1}$$

Enačbo (1) preberemo kot: "Verjetnost, da slučajna spremenljivka X zavzame vrednost k, je μ_k."

Oglejmo si primer, ki modelira prehajanje astronavta med prostori vesoljske ladje.

Zgled 4. *Astronavt sedi na hodniku vesoljske ladje, ki kroži okoli Marsa. Hodnik je sestavljen iz štirih odsekov, ki so krožno povezani. Astronavt vsakih pet minut zamenja odsek, v katerem sedi, ali pa ostane na mestu.*

Delež možnosti, da gre v odsek pred njim, je 0.7*, da se premakne en odsek nazaj* 0.2 *in da ostane v istem odseku* 0.1*. Sprehod začne v prvem odseku. Izjema je drugi odsek, za katerega velja, da je delež možnosti, da gre v odsek pred njim* 0.3*, da se premakne nazaj* 0.3 *in da ostane v istem odseku* 0.4*.*

Množica $S = \{1, 2, 3, 4\}$ je torej množica stanj, sledeče pa prehodna matrika

$$P = \begin{bmatrix} 0.1 & 0.7 & 0 & 0.2 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0.2 & 0.1 & 0.7 \\ 0.7 & 0 & 0.2 & 0.1 \end{bmatrix}.$$

Za neka $i, j \in \{1, 2, 3, 4\}$ element p_{ij}, tj. i-ti element j-tega stolpca, predstavlja verjetnost prehoda iz i-tega v j-ti odsek. Z drugimi besedami, če smo v i-tem odseku, ugotovimo možnost dogodka prehoda iz tega odseka v odsek j tako, da izberemo element na križišču i-te vrstice in j-tega stolpca.

Vidimo, da je verjetnost dogodka različna glede na trenutno stanje in neodvisna od zgodovine stanj, zato lahko proces predstavimo kot Markovsko verigo.

Zanima nas, kakšna je porazdelitev stanja položaja astronavta po treh korakih, torej porazdelitev slučajne spremenljivke X_3. Zaradi preglednosti zapisa bomo to izpeljali na matriki dveh stanj, za večje matrike je rezultat podoben.

Porazdelitev po n korakih dobimo tako, da začetno porazdelitev množimo z n-to potenco prehodne matrike. Oglejmo si to na primeru $n = 3$ za matriko $P \in [0, 1]^{2\times 2}$. Označimo

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

in dobimo

$$P^3 = \begin{bmatrix} p_{11}^{(3)} & p_{12}^{(3)} \\ p_{21}^{(3)} & p_{22}^{(3)} \end{bmatrix},$$

kjer so

$$\begin{aligned}
p_{11}^{(3)} &= p_{11}p_{11}p_{11} + p_{11}p_{12}p_{21} + p_{12}p_{21}p_{11} + p_{12}p_{22}p_{21}, \\
p_{12}^{(3)} &= p_{11}p_{11}p_{12} + p_{12}p_{21}p_{12} + p_{12}p_{22}p_{22} + p_{11}p_{12}p_{22}, \\
p_{21}^{(3)} &= p_{22}p_{22}p_{21} + p_{21}p_{12}p_{21} + p_{21}p_{11}p_{11} + p_{22}p_{21}p_{11}, \\
p_{22}^{(3)} &= p_{22}p_{22}p_{22} + p_{22}p_{21}p_{12} + p_{21}p_{12}p_{22} + p_{21}p_{11}p_{12}.
\end{aligned}$$

Vidimo, da se v vsakem polju matrike nahaja seštevek verjetnosti vseh možnih poti od stanja i do stanja j. Element prve vrstice in prvega stolpca (tj. element $p_{11}^{(3)}$) predstavlja vsoto verjetnosti štirih poti. Prvi seštevanec predstavlja pot, pri kateri trikrat preidemo iz stanja sami vase, tj. $X_0 = 1, X_1 = 1, X_2 = 1, X_3 = 1$, drugi seštevanec zaporedje slučajnih spremenljivk $X_0 = 1, X_1 = 1, X_2 = 2, X_3 = 1$ in ostala seštevanca podobno. Te verjetnosti lahko seštevamo, ker so posamezni izidi – v tem primeru poti dolžine 3 – med seboj disjunktni dogodki.

Zanima nas, kako bi na tej podlagi pridobili odgovor za 3 korake. Če je

$$P = \begin{bmatrix} 0.4 & 0.6 \\ 0.95 & 0.05 \end{bmatrix},$$

dobimo

$$P^3 = \begin{bmatrix} 0.548500 & 0.451500 \\ 0.714875 & 0.285125 \end{bmatrix}.$$

Razberemo lahko, da je verjetnost, da bomo po treh korakih prišli iz stanja 1 nazaj v stanje 1 enaka 0.5485.

Oglejmo si porazdelitev s posebnim imenom.

Definicija 7. ***Stacionarna porazdelitev*** *Markovske verige s prehodno matriko P je takšna porazdelitev π, ki zadošča pogoju*

$$\pi P = \pi.$$

Če je torej trenutno stanje predstavljeno s stacionarno porazdelitvijo, se verjetnostna porazdelitev po enem koraku ne spremeni. Stacionarne porazdelitve imajo pomembno vlogo pri analizi Markovskih verig, najprej pa podajmo še nekaj definicij, ki bodo prostor Markovskih verig dovolj skrčile.

Naj bo M Markovska veriga s prehodno matriko P in množico stanj S. Potem pripadajoči usmerjeni graf G_p definiramo kot

$$G_p = (S, i \to j; P_{ij} > 0).$$

Definicija 8. *Markovska veriga je **periodična**, če je*

$$\gcd(|C|; C \text{ cikel v } G_p) > 1,$$

kjer je gcd *oznaka za največji skupni delitelj (angl.* greatest common divisor*).*

Definicija 9. ***Krepko povezan graf*** *je graf, v katerem obstaja usmerjena pot med dvema poljubnima točkama. Markovska veriga je krepko povezana, če je pripadajoči graf G_p krepko povezan.*

Izrek 1. *Naj bo M Markovska veriga s prehodno matriko P in množico stanj S. Če je M krepko povezana in neperiodična, velja:*

- *M ima enolično stacionarno porazdelitev π,*
- *$\lim_{n\to\infty} \mu P^n = \pi$, kjer je μ poljubna začetna porazdelitev in n število korakov.*

Izrek nam pove, da je verjetnost, da se po dolgo časa znajdemo v stanju i, enaka π_i, pri čemer je π_i i-ti člen stacionarne porazdelitve π. Prav tako je delež preživetega časa v stanju i na dolgi rok enak π_i.

Recimo zdaj še, da je povprečno število korakov med nekima zaporednima obiskoma stanja i enako h_{ii}. Če to "povprečno sliko" skiciramo, vidimo, da smo v k-tem koraku v stanju i, nato sledi $h_{ii} - 1$ stanj različnih od i, nato pa pridemo nazaj v stanje i, potem spet sledi $h_{ii} - 1$ stanj različnih od i in tako naprej. Opazimo, da smo na dolgi rok v $\frac{1}{h_{ii}}$ korakih v stanju i. Velja torej enakost

$$\frac{1}{h_{ii}} = \pi_i.$$

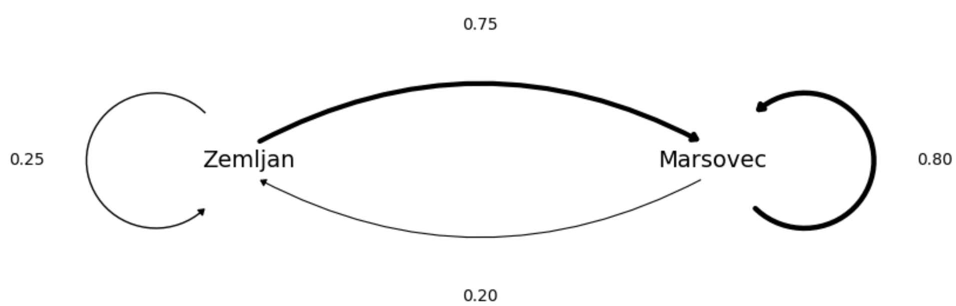
5 MaRSovski problem

Vrnimo se zdaj k začetnemu problemu 1.

Z MaRSovskim problemom smo se spoprijeli z uporabo matrik. Možni stanji pri tem problemu sta Marsovec in Zemljan. Prehodno matriko smo označili s P. V tej matriki p_{11} predstavlja verjetnost, da za Zemljanom stoji Zemljan, p_{12} verjetnost, da za Zemljanom stoji Marsovec, p_{21} verjetnost, da za Marsovcem stoji Zemljan, p_{22} pa verjetnost, da za Marsovcem stoji Marsovec. Dobimo torej

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$$

Markovsko verigo lahko predstavimo tudi z naslednjim grafom:



Slika 2: Predstavitev Markovske verige z uteženim usmerjenim grafom.

V zgledu 4 smo raziskali, kako ugotoviti porazdelitev po končnem številu korakov n. V spodnjem računu so vnešene vrednosti, zastavljene v MaRSovskem problemu. Za $n = 3$ dobimo

$$P = \begin{bmatrix} 0.25 & 0.75 \\ 0.2 & 0.8 \end{bmatrix} \Longrightarrow P^3 = \begin{bmatrix} 0.210625 & 0.789375 \\ 0.210500 & 0.789500 \end{bmatrix}.$$

Vidimo, da je veriga neperiodična, saj vsebuje zanke, je tudi krepko povezana, saj lahko za Zemljanom hodi Marsovec in obratno. Da bi ugotovili delež posameznih vrst osebkov, se lahko poslužimo uporabe izreka. Poiščimo torej stacionarno porazdelitev. Veljati mora naslednji pogoj:

$$[x_1 \ x_2]P = [x_1 \ x_2],$$

kjer vektor $[x_1 \ x_2]$ predstavlja še neznano stacionarno porazdelitev. Pri tem x_1 predstavlja delež Zemljanov, x_2 pa delež Marsovcev.

Dobimo sistem treh enačb z dvema neznankama:

$$\begin{aligned} x_1 &= x_1 p_{11} + x_2 p_{21}, \\ x_2 &= x_1 p_{12} + x_2 p_{22}, \\ 1 &= x_1 + x_2. \end{aligned}$$

Iz druge in tretje enakosti sledi:

$$\begin{aligned} 1 - x_1 &= x_1 p_{12} + (1 - x_1) p_{22} \\ 1 - x_1 &= x_1 p_{12} + p_{22} - x_1 p_{22} \\ 1 - p_{22} &= x_1 + x_1 p_{12} - x_1 p_{22} \\ x_1 &= \frac{1 - p_{22}}{1 + p_{12} - p_{22}}. \end{aligned}$$

Če vstavimo podatke iz MaRSovskega problema, dobimo:

$$\begin{aligned} x_1 &= \frac{1 - p_{22}}{1 + p_{12} - p_{22}} \\ x_1 &= \frac{1 - 0.8}{1 + 0.75 - 0.8} \\ x_1 &= \frac{4}{19}, \quad x_2 = \frac{15}{19}. \end{aligned}$$

Delež Marsovcev na stezi je torej $\frac{15}{19}$, delež Zemljanov pa $\frac{4}{19}$. S tem smo rešili zastavljeni problem.

6 Zaključek

Motivacija za obravnavo Markovskih verig je bil MaRSovski problem, ki ga nismo znali rešiti z nam že znanimi metodami. Spoznali smo torej Markovske verige in pojem stacionarne porazdelitve ter nato zlahka rešili na prvi pogled težek problem.

Literatura

[1] Zapiski s predmeta Verjetnost 2 v šolskem letu 2020/2021 na UL FMF, izvajatelja dr. Janez Bernik in dr. Gregor Šega

[2] M. Williamson, *Markov Chains and Stationary Distributions*. Lane Department of Computer Science and Electrical Engineering, West Virginia University, 2012, dostopno na: `https://community.wvu.edu/~krsubramani/courses/sp12/rand/lecnotes/markov2.pdf`, 29.7.2021

[3] P. Pavešić, *Osnove teorije verjetnosti*, dostopno na: `https://www.fmf.uni-lj.si/~pavesic/POUK/BIOKEMIJA/Matematika\%202\%202009-2010/11.pdf`, 29.7.2021

[4] RadfordMathematics.com, *Probability of "A OR B" : $P(A \cup B)$ Combined Events Part 2*, dostopno na: `https://www.radfordmathematics.com/probabilities-and-statistics/probability-fundamentals/compound-events-probability-A-or-B.html`, 29.7.2021