Markovske verige

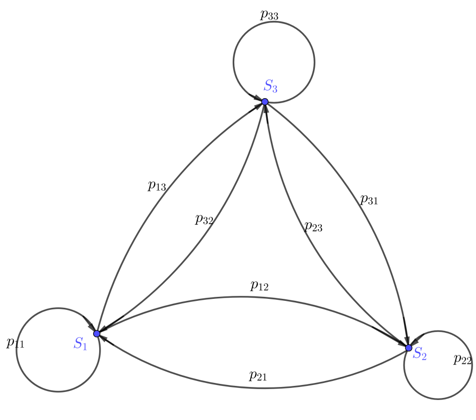
Pod Markovsko verigo si bomo v tem poglavju predstavljali proces, ki se prične v enem izmed *stanj* $S = \{S_1, S_2, \ldots, S_r\}$ in se nato v vsakem naslednjem *koraku* prestavi v neko drugo stanje iz množice S. Če je proces trenutno v stanju S_i, bo v naslednjem koraku v stanju S_j s verjetnostjo p_{ij}. Ta verjetnost je neodvisna od tega, v katerih stanjih je bil proces preden je prišel do stanja S_i. Verjetnost p_{ii} nam pove, s kakšno verjetnostjo ostanemo v stanju S_i. Seveda za vsak $i = 1, 2, \ldots, n$ velja

$$p_{i1} + p_{i2} + \cdots + p_{ir} = 1.$$

Verjetnosti p_{ij} imenujemo *prehodne verjetnosti*, matriko

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

pa *matriko prehodov.* Matrika prehodov ima lastnost, da so vsi elementi nenegativna števila, in da je vsota po vrsticah enaka 1. Taki matriki rečemo *stohastična matrika*, lahko pa tudi *matrika prehodov, verjetnostna matrika, Markovska matrika.*



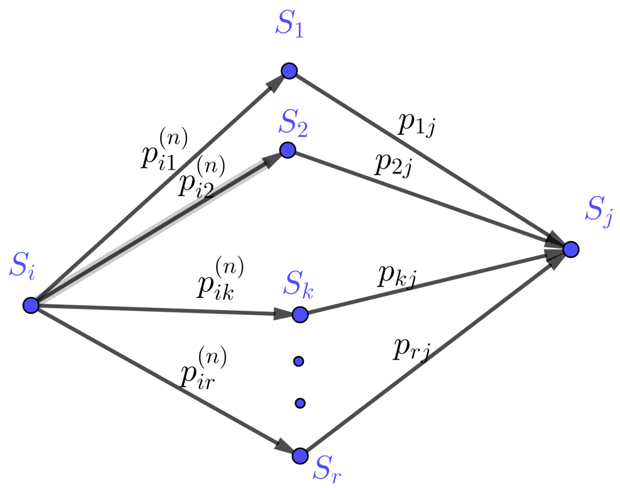
Slika 0.1. Markovska veriga s tremi stanji

Začetno verjetnostno porazdelitev na množici stanj S bomo označili z vodoravnim vektorjem $p = [p_1, p_2, \ldots, p_n]$. Verjetnost p_i nam pove, s kakšno verjetnostjo je začetno stanje procesa ravno stanje S_i. Takemu vektorju rečemo *stohastičen vektor*. Pogosto za začetno stanje kar izberemo eno od stanj, kar pomeni, da je vodoravni vektor p oblike $[0, \ldots, 0, 1, 0, \ldots, 0]$.

TRDITEV. *Naj bo P matrika prehodov za Markovsko verigo, vodoravni vektor p začetna porazdelitev in n naravno število. Velja*

- *(i, j)-ti element matrike P^n nam pove, kakšna je verjetnost, da bo proces po n korakih v stanju S_j, če je trenutno v stanju S_i.*
- *j-ta komponenta vodoravnega vektorja pP^n nam pove verjetnost, da bo proces po n-korakih v stanju S_j, če je začetna porazdelitev stanj enaka p.*

DOKAZ. Označimo z $p_{ij}^{(n)}$ verjetnost, da je proces po n krakih v stanju j, če je trenutno v stanju i. Z indukcijo pokažimo, da je $p_{ij}^{(n)}$ res (i, j)-ti element matrike P^n. Vemo, da velja $p_{ij}^{(1)} = p_{ij}$. Predpostavimo sedaj, da je $p_{ij}^{(n)}$ enak (i, j)-temu elementu matrike P^n in dokažimo analogno izjavo za $n + 1$.

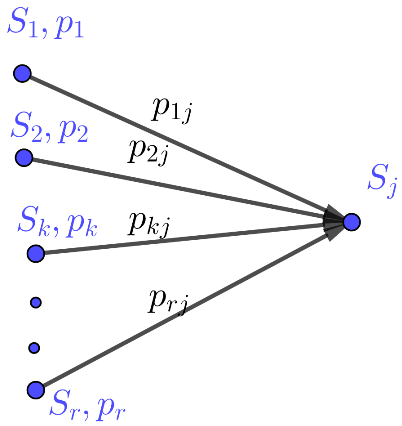


Verjetnost, da je proces po $n + 1$ korakih v stanju S_j, če je trenutno v stanju S_i je enaka

$$p_{ij}^{(n+1)} = p_{i1}^{(n)} p_{1j} + p_{i2}^{(n)} p_{2j} + \cdots + p_{ir}^{(n)} p_{rj},$$

kar pa je ravno (i, j)-ti element matrike $P^n P = P^{n+1}$.

Za drugi del trditve je dovolj, da izjavo pokažemo za $n = 1$, saj po prvi točki potem velja za splošni n.



Verjetnost, da se proces nahaja po prvem koraku v stanju S_j, če je začetna porazdelitev enaka $p = [p_1, p_2, \ldots, p_r]$ je enaka

$$p_1 p_{1j} + p_2 p_{2j} + \cdots + p_2 p_{rj},$$

kar je ravno j-ta komponenta produkta pP. □

Regularne Markovske verige

Za Markovsko verigo rečemo, da je *regularna*, če je možno iz vsakega stanja priti v neko drugo stanje po končno korakih. Če je P matrika prehodov za Markovsko verigo, to pomeni, da mora imeti neka potenca matrike P same neničelne elemente.

Izrek. *Če je P matrika prehodov za Regularno markovsko verigo, potem obstaja stohastičen vektor $w = [w_1, w_2, \ldots, w_r]$, da velja*

$$\lim_{n\to\infty} P^n = W = \begin{bmatrix} w_1 & w_2 & \cdots & w_r \\ w_1 & w_2 & \cdots & w_r \\ \vdots & \vdots & \ddots & \vdots \\ w_1 & w_2 & \cdots & w_r \end{bmatrix}.$$

Vektor w je edini stohastičen vektor, za katerega velja

$$wP = w$$

in za vsak stohastičen vektor p velja

$$\lim_{n\to\infty} pP^n = w.$$

Dokaz. Predpostavimo najprej, da ima že matrika P same neničelne elemente in označimo $q = \min_{i,j} p_{ij}$. Naj bo $x = [x_1, x_2, \ldots, x_r]^T$ poljuben vektor. Označimo $m_0 = \min_k x_k$ in $M_0 = \max_k x_k$. Naj bo $P^n x = [x_1^{(n)}, x_2^{(n)} \ldots, x_r^{(n)}]^T$ in analogno označimo $m_n = \min_k x_k^{(n)}$ in $M_k = \max_k x_k^{(n)}$. Potem velja

$$m_1 = \min_k (p_{k1}x_1 + p_{k2}x_2 + \cdots + p_{kr}x_r) \geq (1-q)m_0 + qM_0 \geq m_0$$

in podobno

$$M_1 = \max_k (p_{k1}x_1 + p_{k2}x_2 + \cdots + p_{kr}x_r) \leq (1-q)M_0 + qm_0 \leq M_0.$$

Če neenakosti odštejemo, dobimo

$$M_1 - m_0 \leq (1-q)M_0 + qm_0 - (1-q)m_0 - qM_0 = (1-2q)(M_0 - m_0).$$

Z indukcijo tako dobimo

$$\begin{aligned} &m_0 \leq m_1 \leq m_2 \leq \cdots \\ &M_0 \geq M_1 \geq M_2 \geq \cdots \\ &(M_n - m_n) \leq (1-2q)^n (M_0 - m_0). \end{aligned}$$

Zaporedji $\{m_n\}$ in $\{M_n\}$ torej konvergirata in imata isto limito. Vektorsko zaporedje $P^n x$ torej konvergira proti vektorju, ki ima same enake elemente. Vzemimo sedaj $x = [0, \ldots, 0, 1, 0, \ldots, 0]^T$, kjer se 1 nahaja na i-tem mestu in naj bo $[w_i, w_i, \ldots, w_i]^T$ limita zaporedja $\{P^n x\}$. Ker je $P^n x$ ravno i-ti stolpec matrike P^n, smo dobili željen rezultat.

V splošnem primeru naj bo N tak, da ima P^N same neničelne elemente in naj bo q minimalna vrednost elementov matrike P^N. Analogno kot zgoraj dobimo

$$M_{N+1} - m_{n+1} \leq (1-2q)(M_N - m_N).$$

Naprej postopamo podobno.

Ker je $\lim_{n\to\infty} P^n = W$, dobimo enakost $WP = W$ in zato $wP = w$. Vektor w je stohastičen, saj je vsota po vrsticah vsake matrike P^n enaka 1 in enakost seveda velja tudi v limiti.

Naj bo sedaj p poljuben stohastičen vektor. Potem velja

$$\lim_{n\to\infty} pP^n = pW = w.$$

S tem smo izrek pokazali. $\square$

Primer. Učitelj ima zgolj tri majice - eno s sliko pujsa, drugo s sliko opice in tretjo s sliko ptiča. Če ima na nek dan oblečeno majico s ptičem, jo bo naslednji dan znova oblekel z verjetnostjo 0.2, vsako izmed ostalih dveh majic pa bo izbral z verjetnostjo 0.4. V primeru, da ima oblečeno majico pujsa ali opice, bo naslednji dan oblekel isto majico z verjetnostjo 0.1, majico s ptičem z verjetnostjo 0.5, tretjo majico pa z verjetnostjo 0.4.

Če po vrsti označimo stanja z S_1 (pujs), S_2 (opica) in S_3 (ptič) je matrika prehodov za markovsko verigo

$$P = \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.4 & 0.1 & 0.5 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}.$$

Ker so vsi elementi matrike neničelni, je matrika regularna. Stohastičen vektor $w = (w_1, w_2, w_3)$, za katerega velja $wP = w$ dobimo tako, da rešimo sistem

$$\begin{aligned} 0.1w_1 + 0.4w_2 + 0.4w_3 &= w_1 \\ 0.4w_1 + 0.1w_2 + 0.4w_3 &= w_2 \\ 0.5w_1 + 0.5w_2 + 0.2w_3 &= w_3. \end{aligned}$$

Če upoštevamo, da mora veljati $w_1 + w_2 + w_3 = 1$ in zato za w_3 vstavimo $1 - w_1 - w_2$, takoj dobimo

$$w_1 = w_2 = \frac{4}{13}, \quad w_3 = \frac{5}{13}.$$

Zgornji izrek nam pove, da matrike P^n konvergirajo proti matriki

$$W = \begin{bmatrix} 4/13 & 4/13 & 5/13 \\ 4/13 & 4/13 & 5/13 \\ 4/13 & 4/13 & 5/13. \end{bmatrix}.$$

Zgornji izrek nam pove, da matrike P^n konvergirajo proti matriki

$$W = \begin{bmatrix} 4/13 & 4/13 & 5/13 \\ 4/13 & 4/13 & 5/13 \\ 4/13 & 4/13 & 5/13 \end{bmatrix} \doteq \begin{bmatrix} 0.31 & 0.31 & 0.38 \\ 0.31 & 0.31 & 0.38 \\ 0.31 & 0.31 & 0.38 \end{bmatrix}.$$

Za ilustracijo konvergence si poglejmo nekaj matrik P^n, pri čemer rezultat zaokrožimo na dve decimalki.

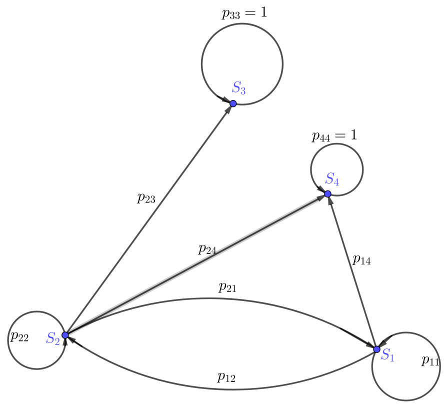
$$P^2 = \begin{bmatrix} 0.37 & 0.28 & 0.35 \\ 0.28 & 0.37 & 0.35 \\ 0.28 & 0.28 & 0.44 \end{bmatrix}$$
$$P^3 = \begin{bmatrix} 0.29 & 0.32 & 0.40 \\ 0.32 & 0.29 & 0.40 \\ 0.32 & 0.32 & 0.37 \end{bmatrix}$$
$$P^4 = \begin{bmatrix} 0.31 & 0.31 & 0.38 \\ 0.31 & 0.31 & 0.38 \\ 0.31 & 0.31 & 0.39 \end{bmatrix}.$$

Prav tako po zgornjem izreku dobimo, da za kateri koli stohastičen vektor p vektorji pP^n konvergirajo proti w. To pomenil, da ne glede na to, katero majico ima na začetni dan oblečeno učitelj, bo imel na dolgi rok majico pujsa in opice bosta oblečeno z verjetnostjo $\frac{4}{13}$, majica ptiča pa s verjetnostjo $\frac{5}{13}$.

Absorbirajoče Markovske verige

Imejmo Markovsko verigo s stanji $S = \{S_1, S_2, \ldots, S_r\}$ in matriko prehodov $P = [p_{ij}]_{i,j=1,\ldots,r}$. Stanje S_k je *absorbirajoče stanje*, če je $p_{kk} = 1$. To seveda pomeni, da so vse ostale prehodne verjetnosti p_{kj}, $j \neq k$, enake 0. Če torej proces pride v stanje S_k, tega stanja ne more več zapustiti.

Markovski proces je *absorbirajoč*, če ima vsaj kakšno absorbirajoče stanje, in če lahko iz vsakega začetnega stanja po končno korakih pridemo v vsaj kakšno absorbirajoče stanje. V absorbirajočem procesu stanja, ki niso absorbirajoča, imenujemo *prehodna stanja*.



SLIKA 0.2. Primer absorbirajoče Markovske verige

Vzemimo sedaj absorbirajočo Markovsko verigo s stanji $S = S_1, S_2, \ldots, S_r$ in predpostavimo, da so $S_1, \ldots, S_m$ absorbirajoča stanja, stanja $S_{m+1}, \ldots, s_r$ pa prehodna stanja. Matrika prehodov P ima torej obliko

$$P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix},$$

kjer je I $m \times m$ identična matrika, 0 $m \times (r-m)$ ničelna matrika, R $(r-m) \times m$ matrika in Q $(r-m) \times (r-m)$ kvadratna matrika. Potenca P^n je potem oblike

$$P^n = \begin{bmatrix} I & 0 \\ R_n & Q^n \end{bmatrix},$$

kjer je R_n primerna $(r-m) \times m$ matrika.

IZREK. *Če imamo absorbirajočo markovsko verigo se vsak markovski proces z verjetnostjo* 1 *zaključi v absorbirajočem stanju.*

DOKAZ. Naj bodo $S_1, \ldots, S_m$ absorbirajoča stanja, $S_{m+1}, \ldots, S_r$ prehodna stanja in

$$P = \begin{bmatrix} I & 0 \\ R & Q \end{bmatrix}$$

matrika prehodov. Za vsak $k \in \{m+1, \ldots, r\}$ naj bo n_k najmanjše število korakov, ki je potrebno, da proces, ki se prične v S_k lahko doseže kakšno absorbirajoče stanje. Če je $n = \max_k n_k$, lahko v n korakih iz katerega koli stanja pridemo v neko absorbirajoče stanje. Če je torej v n-tem koraku proces v enem od prehodnih stanj, je verjetnost, da je v $n+1$ koraku proces v enem od prehodnih stanj strogo manjši od 1. Označimo to verjetnost s q. Zato je verjetnost, da je proces v enem od prehodnih stanj v $(n+l)$-tem koraku manjša ali enaka q^l. Ker je $q < 1$, gre verjetnost, da je proces v enem od absorbirajočih stanj, proti 1. □

Zgornji izrek je ekvivalenten izjavi, da gredo potence Q^n proti ničelni matriki.

TRDITEV. *Matrika $I - Q$ je obrnljiva in velja*

$$(I - Q)^{-1} = I + Q + Q^2 + Q^3 + \cdots$$

DOKAZ. Za obrnljivost moramo pokazati le, da 1 ni lastna vrednost matrike Q. Recimo, da je $Qx = x$. Potem velja $Q^n x = x$ in zato

$$x = \lim_{n\to\infty} Q^n x = 0.$$

Označimo sedaj N inverz matrike $I - Q$. Velja

$$(I - Q)(I + Q + Q^2 + \ldots + Q^n) = I - Q^{n+1}$$

in zato

$$I + Q + Q^2 + \ldots + Q^n = N(I - Q^{n+1}) = N - NQ^{n+1}.$$

V limiti dobimo

$$I + Q + Q^2 + \ldots = N.$$

□

Matriko $N = (I - Q)^{-1}$ imenujemo *fundamentalna matrika* absorbirajočega Markovske verige.

IZREK. *Naj bo $N = (I - Q)^{-1} = [n_{ij}]$. Potem je n_{ij} pričakovano število nahajanj procesa v prehodnem stanju S_{m+j}, če se proces prične v stanju S_{m+i}.*

DOKAZ. Predpostavimo, da se proces prične v stanju S_{m+i} in fiksirajmo j. Naj bo X_k slučajna spremenljivka, ki je enaka 1, če je v k tem koraku proces v stanju S_{m+j} in 0, če je v katerem drugem stanju. Pričakovana vrednost nahajanj v stanju S_{m+j} do n-tega koraka je enaka

$$E(X_0 + X_1 + \ldots + X_n) = q_{ij}^{(0)} + q_{ij}^{(1)} + \ldots + q_{ij}^{(n)},$$

kjer z $q_{ij}^{(k)}$ označimo (i, j)-ti element matrike Q^k. V limiti torej dobimo ravno

$$E(X_0 + X_1 + \ldots) = n_{ij}.$$

□

POSLEDICA. *Pričakovano število potrebnih korakov, da se Markovski proces, ki se prične v prehodnem stanju S_{m+k}, zaključi a nekem absorbirajočem stanju, je enako vsoti elementov k-te vrstice matrike N.*

Označimo sedaj še

$$B = [b_{ij}] = NR.$$

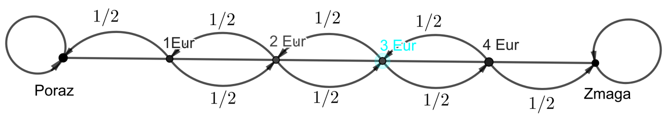
IZREK. *Verjetnost, da se Markovski proces, ki se prične v prehodnem stanju S_{m+k} zaključi v absorbirajočem stanju S_j, je enaka b_{kj}.*

DOKAZ. Verjetnost, da se proces, ki se je pričel v stanju S_{m+i} zaključi v stanju S_j je enaka

$$\sum_{n=0}^{\infty} \sum_{k=1}^{m-r} q_{ik}^{(n)} r_{kj} = \sum_{k=1}^{m-r} \sum_{n=0}^{\infty} q_{ik}^{(n)} r_{kj} = \sum_{k=1} n_{ik} r_{kj} = b_{ij}.$$

□

PRIMER (Kockarjev propad). Dva igralca mečeta kovanec. Če pade cifra, prvi igralec plača drugemu en evro, če pade grb, pa drugi igralec prvemu plača en evro. Predpostavimo, da imamo pošten kovanec in da ima prvi igralec v začetku tri evre, drugi pa dva evro. Igra se konča, ko eden od igralcev izgubi ves svoj denar.



SLIKA 0.3. Kockarjev propad

Matrika prehodov s stališča prvega igralca, pri čemer smo obe stanja pisali v vrstnem redu: Poraz, Zmaga, 1 Eur, 2 Eur, 3 Eur, 4 Eur, je enaka

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 & 0 & 1/2 & 0 \end{bmatrix}.$$

Z malo truda dobimo

$$N = (I-Q)^{-1} = \begin{bmatrix} 1 & -1/2 & 0 & 0 \\ -1/2 & 1 & -1/2 & 0 \\ 0 & -1/2 & 1 & -1/2 \\ 0 & 0 & -1/2 & 1 \end{bmatrix}^{-1} = \frac{1}{5}\begin{bmatrix} 8 & 6 & 4 & 2 \\ 6 & 12 & 8 & 4 \\ 4 & 8 & 12 & 6 \\ 2 & 4 & 6 & 8 \end{bmatrix}$$

in

$$B = NR = \frac{1}{5}\begin{bmatrix} 8 & 6 & 4 & 2 \\ 6 & 12 & 8 & 4 \\ 4 & 8 & 12 & 6 \\ 2 & 4 & 6 & 8 \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1/2 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 4 & 1 \\ 3 & 2 \\ 2 & 3 \\ 1 & 4 \end{bmatrix}.$$

Pričakovano število metov kovanca, da se igra konča, je $6 = (4+8+12+6)/5$, prvi igralec pa bo igro dobil z verjetnostjo $3/5$ in izgubil z verjetnostjo $2/5$.