Sveučilište u Zagrebu

FAKULTET STROJARSTVA I BRODOGRADNJE

Podloge za vježbe iz kolegija

Umjetna inteligencija

Markovljevi lanci

Bojan Šekoranja

1. Markovljevi lanci – teorijska osnova

1.1. Vjerojatnost

Pod slučajnim pokusom podrazumijevamo takav pokus čiji ishodi, tj. rezultati nisu jednoznačno određeni uvjetima u kojima izvodimo pokus. Mogućnost ponavljanja danog slučajnog pokusa je proizvoljno konačno mnogo puta. Elementarni događaj

$$\omega_1, \omega_2, \dots \omega_n \tag{1.1}$$

je svaki od konačno mnogo ishoda slučajnog pokusa. Osnovni polazni objekt u teoriji vjerojatnosti jest neprazan skup

$$\Omega = \{\omega_1, \omega_2, \dots \omega_n\}, \tag{1.2}$$

koji zovemo prostorom elementarnih događaja i on reprezentira skup svih ishoda slučajnog pokusa. Cijeli prostor Ω zovemo sigurnim događajem, a prazan skup $\emptyset$ nemogućim događajem. Familija $\mathcal{F}$ podskupova od Ω je σ-algebra skupova na Ω ako vrijedi:

1. $\Omega \in \mathcal{F}, \emptyset \in \mathcal{F}$, (1.3)
2. $\mathrm{A} \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$, (1.4)
3. $\mathrm{A}, \mathrm{B} \in \mathcal{F} \Rightarrow \mathrm{A} \cup \mathrm{B} \in \mathcal{F}$. (1.5)

Vjerojatnost je preslikavanje

$$P: \mathcal{F} \to [0,1] \tag{1.6}$$

koje, ako vrijede prethodna svojstva, zajedno s Ω i $\mathcal{F}$ čini vjerojatnosni prostor

$$(\Omega, \mathcal{F}, \mathrm{P}). \tag{1.7}$$

Neka su A i B događaji, $\mathrm{A}, \mathrm{B} \subset \Omega$. Događaj koji se ostvaruje ako se ostvario barem jedan od događaja A, B naziva se unija događaja i označava s $\mathrm{A} \cup \mathrm{B}$. Događaj koji se ostvaruje ako su se ostvarila oba događaja A i B naziva se presjek događaja i označava s $\mathrm{A} \cap \mathrm{B}$. Za događaje A,B skup A\B, koji nazivamo razlika događaja, je događaj koji nastane ako nastane događaj A, a ne nastane događaj B.

Vjerojatnost unije događaja A i B, $\mathrm{A}, \mathrm{B} \subset \Omega$

$$\mathrm{P}(\mathrm{A} \cup \mathrm{B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B}) - \mathrm{P}(\mathrm{A} \cap \mathrm{B}). \tag{1.8}$$

1.2. Uvjetna vjerojatnost

Neka su A i B događaji iz vjerojatnosnog prostora $(\Omega, \mathcal{F}, \mathrm{P})$. Računamo vjerojatnost da se dogodio događaj A uz uvjet da nam je poznato da se dogodio događaj B. U tom je slučaju

dovoljno samo promotriti elementarne događaje koji čine B i među njima tražiti one povoljne za događaj A. U tom slučaju tražimo elementarne događaje za presjek $A \cap B$ tih događaja. Ova vjerojatnost ovisi o događaju B i naziva se uvjetna vjerojatnost:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.9}$$

Događaji A i B su nezavisni ako informacija o realizaciji događaja B ne utječe na vjerojatnost događaja A i vrijedi bilo koja od jednakosti: $P(A|B) = P(A)$ ili $P(A|B) = P(B)$. Iz toga slijedi:

$$P(A|B) = P(A) = \frac{P(A \cap B)}{P(B)}, \tag{1.10}$$

$$P(A \cap B) = P(A) \cdot P(B). \tag{1.11}$$

1.3. Stohastički procesi

Slučajna varijabla na vjerojatnosnom prostoru $(\Omega,\ \mathcal{F}, P)$ je preslikavanje

$$X: \Omega \to \mathbb{R}. \tag{1.12}$$

Slučajna varijabla neovisna je o vremenu. Međutim, mnogi procesi čiji je ishod neizvjestan, a koji se odvijaju u vremenu zahtijevaju da se koncept slučajne varijable poopći tako da uključuje i vremensku komponentu. Neka je $T \subset \mathbb{R}$ skup vremena u kojemu promatramo stohastički proces. Za svako vrijeme $t \in T$ određena je slučajna varijabla koju ćemo označavati s X_t.

Uvodimo definiciju:

Stohastički (slučajni) proces je skup slučajnih varijabli

$$X = \{X_t, t \in T\}. \tag{1.13}$$

T je parametarski skup stohastičkog procesa, a $t \in T$ je parametar.

Stohastički proces možemo shvatiti i kao funkciju dviju varijabli, skupa vremena T i skupa stanja S (skup unutar kojeg proces poprima vrijednosti)

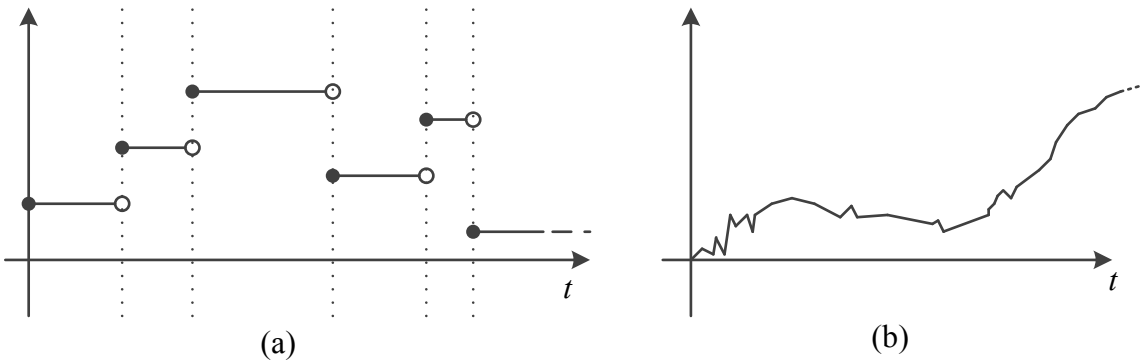
$$X: T \times \Omega \to S. \tag{1.14}$$

Stohastički procesi razlikuju se po prirodi skupa T u dvije grupe: procesi kontinuirani u vremenu i diskretni procesi. Ako je skup T interval, govorimo o kontinuiranom stohastičkom procesu, a u slučaju da je skup T prebrojiv tj. $T = \{t_0,\ t_1,\ t_2, \dots\}$ govorimo o diskretnom stohastičkom procesu. Pritom i slučajne varijable stohastičkog procesa (skup stanja S) mogu biti diskretne ili neprekidne.

Tablica 1.1. *Podjela stohastičkih procesa.*

		Skup T	
		Diskretan	Kontinuiran
Skup S	Diskretan	Markovljevi lanci	Poissonov proces
	Kontinuiran	-	Brownovo gibanje

Teorija Markovljevih lanaca proučava nizove slučajnih varijabli kod kojih je su skup stanja T i skup stanja S diskretni. Poissonov proces (Slika *1.1* a) primjer je procesa s kontinuiranim vremenom T i diskretnim skupom stanja S. Drugi primjer procesa s kontinuiranim vremenom T, ali s kontinuiranim skupom stanja S je Brownovo gibanje (Slika *1.1* b). Trajektorija jednodimenzionalnog Brownova gibanja neprekinuta je funkcija.



Slika 1.1. *Stohastički procesi: a) Poissonov proces, b) Brownovo gibanje.*

Stohastički procesi mogu se podijeliti na stacionarne i nestacionarne. Stacionarni su oni procesi kod kojih su vjerojatnosne osobine invarijantne u odnosu na pomake vremenskog parametra, dok se kod nestacionarnih procesa te osobine mogu mijenjati.

1.4. Markovljevi lanci

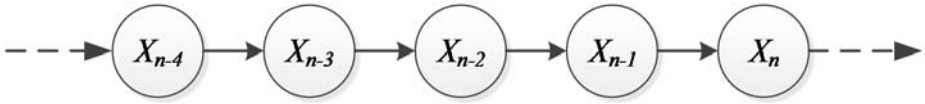
Stohastičke procese koji imaju tzv. „svojstvo zaboravljivosti" nazivamo Markovljevim lancima. Markovljev lanac, nazvan po Andreyu Markovu, predstavlja niz stanja sustava. U svakom trenutku sustav može prijeći u neko novo stanje ili može ostati u istom stanju. Promjene stanja nazivaju se tranzicije.

Niz diskretnih slučajnih varijabli $X_0, X_1, \ldots$. zvat ćemo stohastičkim lancem. Slučajne varijable uzimaju vrijednosti u konačnom skupu $S = \{s_0,\ s_1, \ldots,\ s_n\}$.

Lanac X_0, X_1, … je Markovljev lanac prvog reda (Slika *1.2*), ako za sve izbore stanja $s_0, s_1, \ldots, s_n \in S$ vrijedi:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1}) \qquad (1.15)$$

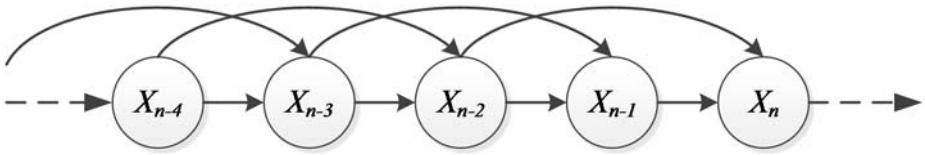
Ovdje trenutak *n* predstavlja sadašnjost, a *0,..., n-1* prošlost. Sadašnje stanje s_n ovisi samo o prethodnom s_{n-1}, ali ne i o načinu na koji je proces dospio u prethodno stanje tj. vrijednostima procesa u ranijim trenutcima.



Slika 1.2. *Markovljev lanac prvog reda.*

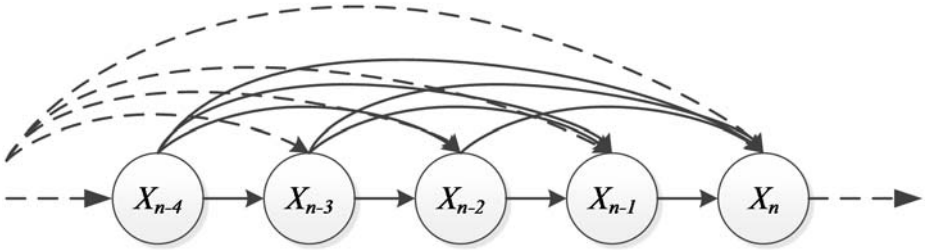
Markovljev lanac drugog reda (Slika *1.3*) ovisi o dvama prethodnim stanjima s_{n-1}, s_{n-2} te vrijedi:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2}) \qquad (1.16)$$



Slika 1.3. *Markovljev lanac drugog reda.*

Sukladno tome, Markovljev lanac *k*-tog reda (Slika *1.4*) ovisi o *k* prethodnih stanja s_{n-1}, $s_{n-2}, \ldots s_{n-k}$:



Slika 1.4. *Markovljev lanac k-tog reda.*

Viši redovi lanaca „pamte” dalje u prošlost tj. daju više informacija o procesu.

1.5. Prijelazne vjerojatnosti

Markovljeve lance možemo podijeliti na stacionarne i nestacionarne. Za Markovljeve lance koji imaju svojstvo stacionarnosti prijelazne vjerojatnosti ne ovise o koraku, odnosno trenutku. Veza između slučajnih varijabli X_n i X_{n-1} zadana je prijelaznim vjerojatnostima. Vjerojatnost prijelaza iz stanja s_i u stanje s_j je p_{ij}

$$p_{ij} = P\left(X_n = s_j \middle| X_{n-1} = s_i\right). \tag{1.17}$$

Matrica s elementima p_{ij} označava se s P i naziva se matrica prijelaznih vjerojatnosti.

$$P = (p_{ij})\ i, j \in \{1,2,\ldots,k\}. \tag{1.18}$$

Matrica prijelaznih vjerojatnosti prvog reda za k - broj stanja

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{pmatrix}. \tag{1.19}$$

Elementi ove matrice su nenegativni, $p_{ij} \geq 0$, a zbroj elemenata u svakom njezinom retku jednak je jedan:

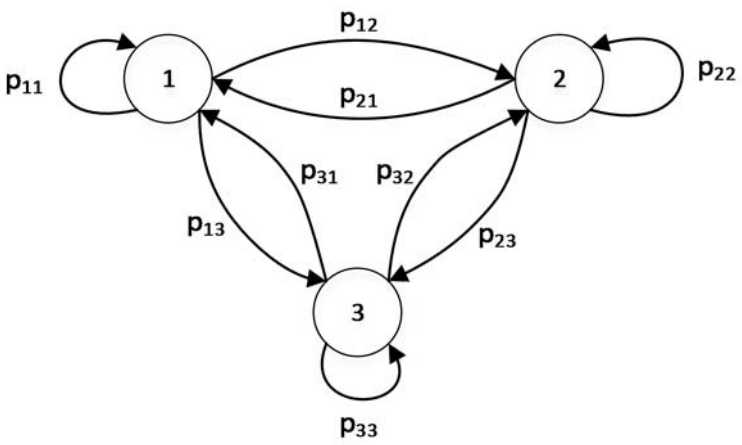
$$p_{ij} \geq 0, \quad \sum_{j=1}^{k} p_{ij} = 1 \quad \forall i, j \in \{1,2,\ldots,k\}. \tag{1.20}$$

Primjerice matrica prijelaznih vjerojatnosti P Markovljeva lanca s tri stanja {1, 2, 3} (Slika *1.5*) glasi

$$\mathrm{P} = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}, \tag{1.21}$$

pri čemu je

$$\sum_{j=1}^{3} p_{ij} = 1, i = 1,2,3. \tag{1.22}$$



Slika 1.5. *Markovljev lanac s tri stanja.*

U slučaju nestacionarnih lanaca potrebno je definirati matricu prijelaznih vjerojatnosti u n-tom koraku

$$P(n) = \left(p_{ij}^{(n)}\right). \qquad (1.23)$$

2. Primjeri zadataka

2.1. Zadatak 1

Prikazati grafom zadane matrice prijelaznih vrijednosti

a) $\boldsymbol{P} = \begin{bmatrix} 0{,}5 & 0{,}5 \\ 0{,}5 & 0{,}5 \end{bmatrix}$

b) $\boldsymbol{P} = \begin{bmatrix} 0{,}5 & 0{,}5 \\ 1 & 0 \end{bmatrix}$

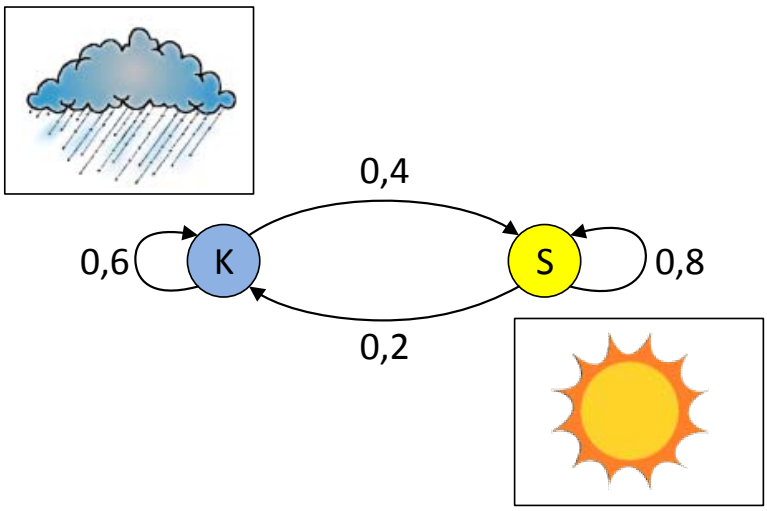
c) $\boldsymbol{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

d) $\boldsymbol{P} = \begin{bmatrix} 1/3 & 0 & 2/3 \\ 0 & 1 & 0 \\ 0 & 2/3 & 1/3 \end{bmatrix}$

e) $\boldsymbol{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$

2.2. Zadatak 2

Zadan je graf dva stanja s prijelaznim vjerojatnostima



a) oblikovati matricu prijelaznih vjerojatnosti
b) izračunati vjerojatnosti za stanje $\boldsymbol{X}_1$ ako je u početnom stanju $\boldsymbol{X}_0$ sunce
c) izračunati vjerojatnosti za stanje $\boldsymbol{X}_3$ ako je u početnom stanju $\boldsymbol{X}_0$ kiša
d) izračunati vjerojatnosti za stanje $\boldsymbol{X}_\infty$

a) Matrica prijelaznih vjerojatnosti **P**

$$\boldsymbol{P} = \begin{array}{c c} & \begin{array}{cc} S & K \end{array} \\ \begin{array}{c} S \\ K \end{array} & \begin{bmatrix} 0{,}8 & 0{,}2 \\ 0{,}4 & 0{,}6 \end{bmatrix} \end{array}$$

$$\boldsymbol{P} = \begin{bmatrix} 0{,}8 & 0{,}2 \\ 0{,}4 & 0{,}6 \end{bmatrix}$$

b) $\boldsymbol{X}_1 =?$

$$\boldsymbol{X}_0 = \begin{array}{c} \begin{array}{cc} S & K \end{array} \\ \begin{bmatrix} 1 & 0 \end{bmatrix} \end{array}$$

$$\boldsymbol{X}_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\boldsymbol{X}_1 = \boldsymbol{X}_0\boldsymbol{P} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0{,}8 & 0{,}2 \\ 0{,}4 & 0{,}6 \end{bmatrix}$$

$$\boldsymbol{X}_1 = \begin{bmatrix} 0{,}8 & 0,2 \end{bmatrix}$$

c) $\boldsymbol{X}_3 =?$

$$\boldsymbol{X}_3 = \boldsymbol{X}_2\boldsymbol{P} = (\boldsymbol{X}_1\boldsymbol{P})\boldsymbol{P} = \boldsymbol{X}_1\boldsymbol{P}^2 = \boldsymbol{X}_0\boldsymbol{P}^3$$

$$\boldsymbol{P}^2 = \boldsymbol{P}\boldsymbol{P} = \begin{bmatrix} 0.720 & 0.280 \\ 0.560 & 0.440 \end{bmatrix}$$

$$\boldsymbol{P}^3 = \boldsymbol{P}^2\boldsymbol{P} = \begin{bmatrix} 0.688 & 0.312 \\ 0.624 & 0.376 \end{bmatrix}$$

$$\boldsymbol{X}_3 = \boldsymbol{X}_0\boldsymbol{P}^3 = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.688 & 0.312 \\ 0.624 & 0.376 \end{bmatrix}$$

$$\boldsymbol{X}_3 = \begin{bmatrix} 0.624 & 0{,}376 \end{bmatrix}$$

d) $\boldsymbol{X}_\infty =?$

Ako broj stanja iznosi dva i prijelazne vjerojatnosti su stacionarne moguć je izračun vjerojatnosti za $\boldsymbol{X}_\infty$ uz pomoć rješavanja jednadžbi. U suprotnome ako se radi o većem broju stanja potrebno je drugim metodama utvrditi $\boldsymbol{P}^\infty$.

$$P(X_\infty = S) = P(S|S)P(X_\infty = S) + P(S|K)P(X_\infty = K)$$
$$P(X_\infty = K) = P(K|S)P(X_\infty = S) + P(K|K)P(X_\infty = K)$$

$$P(X_\infty = S) = 0{,}8P(X_\infty = S) + 0{,}4P(X_\infty = K)$$
$$P(X_\infty = K) = 0{,}2P(X_\infty = S) + 0{,}6P(X_\infty = K)$$

$$0{,}2P(X_\infty = S) = 0{,}4P(X_\infty = K)$$
$$0{,}4P(X_\infty = K) = 0{,}2P(X_\infty = S)$$

$$P(X_\infty = S) = 2P(X_\infty = K)$$
$$P(X_\infty = K) = 0{,}5P(X_\infty = S)$$

Vjerojatnost uvijek iznosi 1 te možemo postaviti jednadžbu

$$P(X_\infty = S) + P(X_\infty = K) = 1$$

$$\Downarrow$$

$$P(X_\infty = S) = \frac{2}{3}$$
$$P(X_\infty = K) = \frac{1}{3}$$

$$\Downarrow$$

$$\boldsymbol{X}_\infty = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

2.3. Zadatak 3

Pretpostavimo da se osobe po struci može klasificirati kao matematičar, fizičar i strojar. Pretpostavimo da je kod potomaka strojara 80% strojara, 10% matematičara i 10% su fizičari. U slučaju fizičara 50% je fizičara, 20% matematičara i 30% strojara. Konačno, u slučaju matematičara 50% potomaka su matematičari, a po 25% u druge dvije kategorije. Pretpostavimo da svaka osoba ima barem jednog potomka i oblikuje Markovljev lanac slijedeći profesiju slučajno odabranog potomka dane obitelji kroz nekoliko generacija.

a) Postavite matricu prijelaznih vjerojatnosti.

b) Pronađite vjerojatnost da je slučajno izabrani unuk fizičara strojar

Rješenje:

a) $S = \{S, M, F\} \rightarrow$ P

b) P= 0,33

2.4. Zadatak 4

Imam 4 kišobrana, neke kod kuće, neke u uredu. Krećem se između te dvije lokacije. Kišobran nosim sa sobom samo kada pada kiša. Ako ne pada kiša ne ponesem kišobran. Postoji mogućnost da su sva 4 kišobran na jednom od dva moguća mjesta, a ja se nalazim na drugom – tamo gdje nema kišobrana. Trenutna procjena pokazuje da je vjerojatnost kiše $\alpha = 0{,}6$ u Zagrebu.

a). Ako je vjerojatnost kiše α, koja je vjerojatnost da pokisnem?

b) Koliko kišobrana trebam imati tako da, ako slijedim gore navedenu strategiju, vjerojatnost da pokisnem bude manja od 0,01?

Rješenje:

a) P(pokisnem)≈0,0545
b) N=23,6 – potrebno je 24 kišobrana

3. Dodatno

3.1. Skriveni Markovljevi lanci – HMM

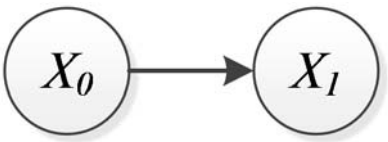
Kod markovljevih lanaca imali smo slučaj da je proces vidljiv tj da u svakom trenu znamo stanje sustava tj na našem primjeru da li vani pada kiše ili je sunce. U slučaju skrivenog markovljevog lanca (HMM – Hidden Markov Model) nemamo informaciju o stanju sustava, već imamo samo opservacije. HMM je MM sa skrivenom sekvencom stanja. To je dvostruki stohastički proces koji se sastoji od skrivenog stohastičkog proces (sekvenca stanja) i vidljivog stohastičkog procesa koji generira sekvencu opažanja.

$$P(X_1, E_1, \ldots, X_T, E_T) = P(X_1)P(E_1|X_1)\prod_{t=2}^{T} P(X_t|X_{t-1})P(E_t|X_t)$$

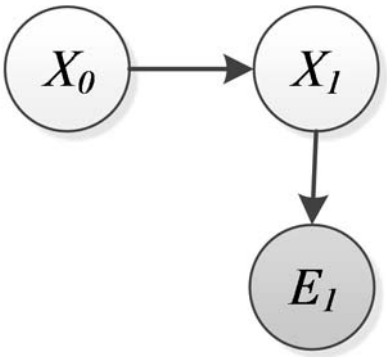
Prethodno stanje — Prijelazni model — Opservacijski model

Estimacija stanja (filtriranje)

Praćenje distribucije $B_t(X)=P_t(X_t \mid e_1,\ldots,e_t)$ (vjerovanje) kroz vrijeme. Počinjemo s $B(X)$ koji je postavljen inicijalno, obično uniformno. Prolaskom vremena dobivamo nova opažanja i u skladu s njima ažuriramo $B(X)$.



$$B'(X_{t+1}) = \sum_{x_t} P(X'|x_t)\, B(x_t)$$



$$B(X) \propto P(e|X)B'(X)$$

3.2. Primjer zadatka

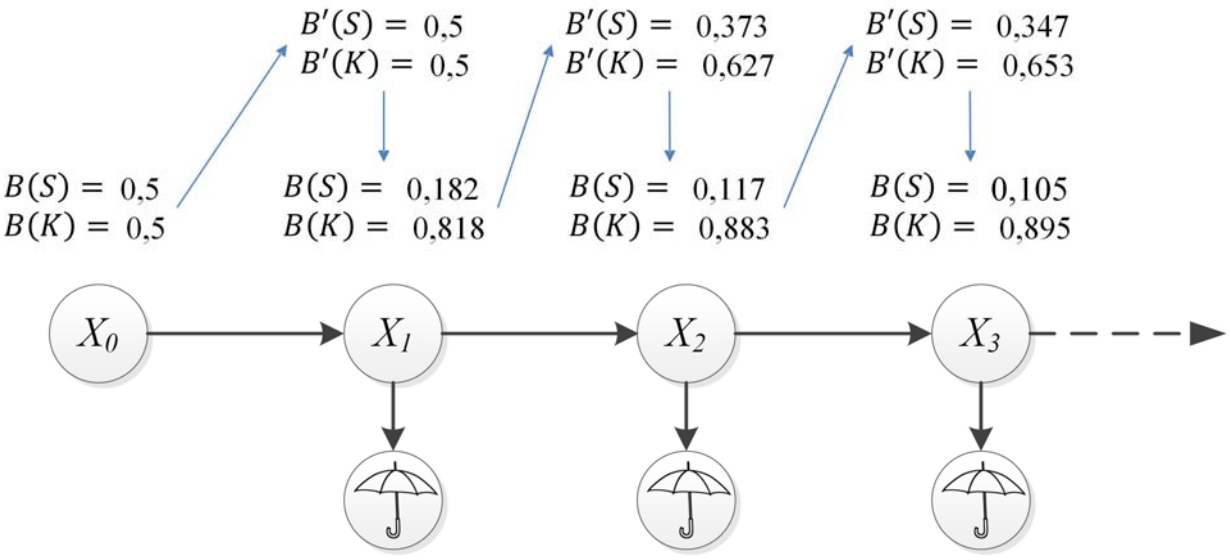
Učite za ispit zatvoreni danima u zgradi bez prozora te nemate informaciju o tome kakvo je vrijeme vani. Jedino što znate je kada vidite kolegu koji prolazi po hodniku i da li on nosi kišobran ili ne. Naravno ni time ne možete sigurno znati jer kolega isto tako nekad ne ponese kišobran i pokisne, a nekada ga ponese iako je sunčano.

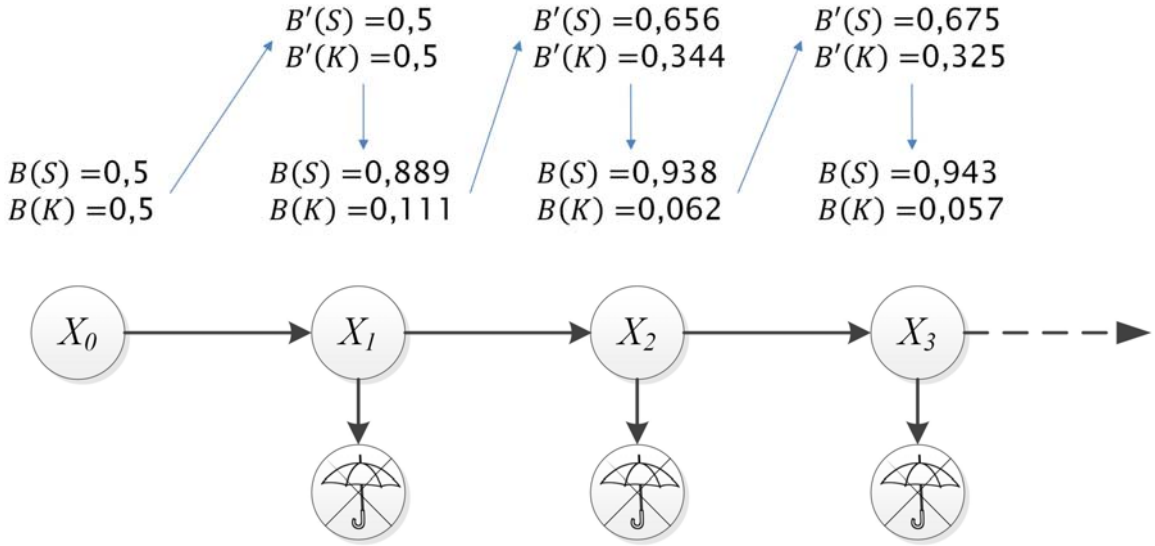
Poznate su prijelazne vrijednost između sunčanih i kišnih dana:

X_{t-1}	X_t	
	sunce	kiša
sunce	0,7	0,3
kiša	0,3	0,7

Poznate su i vjerojatnosti opservacije s obzirom na određeno stanje:

	sunce	kiša
kišobran	0,2	0,9
¬kišobran	0,8	0,1


B′(S) = 0,5
B′(K) = 0,5
B′(S) = 0,373
B′(K) = 0,627
B′(S) = 0,347
B′(K) = 0,653
B(S) = 0,5
B(K) = 0,5
B(S) = 0,182
B(K) = 0,818
B(S) = 0,117
B(K) = 0,883
B(S) = 0,105
B(K) = 0,895
X0
X1
X2
X3


B′(S) =0,5
B′(K) =0,5
B′(S) =0,656
B′(K) =0,344
B′(S) =0,675
B′(K) =0,325
B(S) =0,5
B(K) =0,5
B(S) =0,889
B(K) =0,111
B(S) =0,938
B(K) =0,062
B(S) =0,943
B(K) =0,057
X0
X1
X2
X3