

Megszámlálhatóan végtelen állapotterű Markov-láncok

1 Bevezetés

Olyan Markov-láncokról lesz szó, melyek állapottere valamely megszámlálhatóan végtelen halmaz. Ilyen halmazok például az alábbiak:

- természetes számok: $\mathbb{N}$,
- egész számok: $\mathbb{Z}$,
- a sík azon pontjai, melyeknek koordinátái egész számok (vagyis a rácspontok): $\mathbb{Z}^2$,
- a d dimenziós tér azon pontjai, melyeknek koordinátái egész számok: $\mathbb{Z}^d$,
- racionális számok: $\mathbb{Q}$,
- stb.

A fenti halmazok és a természetes számok halmaza között egy-egy értelmű megfeleltetés létesíthető. Ezzel szemben a következő halmazok nem megszámlálhatóan végtelen számosságúak („sokkal több elemük van, mint $\mathbb{N}$ -nek”).

- valós számok: $\mathbb{R}$,
- irracionális számok: $\mathbb{R} \setminus \mathbb{Q}$,
- komplex számok: $\mathbb{C}$,

- a sík pontjai: $\mathbb{R}^2$,
- a d dimenziós tér pontjai: $\mathbb{R}^d$,
- a $[0, 1]$ intervallum pontjai,
- stb.

Az előzőekhez hasonlóan jelölje most is X_n a Markov-láncot, S pedig a lehetséges állapotok halmazát, vagyis az állapotteret. A véges állapotterű Markov-láncoknál megismert tulajdonságok közül számos most is érvényes lesz, néhány esetleg egy kicsit más formában. Itt is beszélhetünk átmenetvalószínűség mátrixról, de ez most egy végtelen mátrix lenne, ezért nem használjuk a mátrixos írásmódot, hanem csak egyszerűen az átmenetvalószínűségeket:

$$p(x, y) = P(X_1 = y | X_0 = x), \quad x, y \in S.$$

Az átmenetvalószínűségek nyilvánvalóan nem negatívak, és minden $x \in S$ esetén teljesül, hogy

$$\sum_{y \in S} p(x, y) = 1.$$

Beszélhetünk az n -lépéses átmenetvalószínűségekről is:

$$p_n(x, y) = P(X_n = y | X_0 = x).$$

Tekintsük az állapottér x, y, z elemeit. Ekkor $0 < m, n < \infty$ esetén

$$\begin{aligned} p_{m+n}(x, y) &= P(X_{m+n} = y | X_0 = x) = \sum_{z \in S} P(X_{m+n} = y, X_m = z | X_0 = x) \\ &= \sum_{z \in S} p_m(x, z) p_n(z, y). \end{aligned}$$

Ezt az egyenletet CHAPMAN-KOLMOGOROV-egyenletnek nevezik.

Példák

1. **Bolyongás félig elnyelő fallal:** Legyen az állapottér $S = \{0, 1, 2, \dots\}$ és $0 < p < 1$. Az átmenetvalószínűségek:

$$p(x, x-1) = 1-p, \quad p(x, x+1) = p, \quad \text{ha } x > 0, \quad \text{és} \quad p(0, 0) = 1-p, \quad p(0, 1) = p.$$

2. **Szimmetrikus véletlen bolyongás $\mathbb{Z}^d$ -ben:** Jelölje $\mathbb{Z}^d$ a d dimenziós tér azon pontjainak halmazát, melyeknek minden koordinátája egész szám. Figyeljük meg, hogy $\mathbb{Z}^d$ minden elemének pontosan $2d$ szomszédja van (olyan elemek, melyek tőle egységnyi távolságra vannak). A véletlen bolyongás során minden egyes lépésben csak szomszédos pontba ugorhatunk át, egyenlő valószínűséggel választva a lehetőségek közül. Tehát egy olyan Markov-lánc írja le a folyamatot, melynek állapottere $S = \mathbb{Z}^d$, az átmenetvalószínűségek pedig

$$p(x, y) = \begin{cases} \frac{1}{2d}, & \text{ha } |x - y| = 1, \\ 0, & \text{különben.} \end{cases}$$

3. **Sorbanállási modell:** Tekintsünk egy olyan tömegkiszolgálási rendszert, amelyben egyetlen kiszolgálóegység van, a várakozó igények alkotta sor hossza pedig nem korlátozott. Minden egyes időintervallumban p valószínűséggel érkezik új igény, a legelső igényt pedig q valószínűséggel elégítik ki (ekkor ez az igény távozik a rendszerből). Legyen X_n a rendszerben levő igények száma (az éppen kiszolgálás alatt levővel együtt). Ekkor ez a folyamat Markov-lánc, melynek állapottere $S = \{0, 1, 2, \dots\}$, az átmenetvalószínűségek pedig

ha $x > 0$, akkor:

$$p(x, x+1) = p(1-q), \quad p(x, x-1) = (1-p)q, \quad p(x, x) = pq + (1-p)(1-q)$$

ha $x = 0$, akkor:

$$p(0, 0) = 1-p, \quad p(0, 1) = p.$$

Miként a véges állapotterű Markov-láncoknál, most is a hosszú távú viselkedés vizsgálata a célunk. Itt is beszélhetünk kapcsolódó állapotokról, osztályokról. Ha csak egyetlen osztály létezik, akkor a Markov-láncot most is irreducibilisnek nevezzük. A periódus fogalma is ugyanúgy értelmezett, mint véges esetben. Azonban elmentésben a véges állapotterű esettel, itt nem mindig igaz, hogy irreducibilis, aperiódikus Markov-lánc az invariáns eloszláshoz tart.

2 Visszatérőség

Legyen X_n egy irreducibilis Markov-lánc megszámlálhatóan végtelen S állapottérrel. Azt mondjuk, hogy X_n *rekurrens*, ha minden $x \in S$ állapotra igaz, hogy

$$P(X_n = x \text{ végtelen sok } n\text{-re}) = 1,$$

vagyis a lánc végtelen sokszor visszatér az x állapotba. Ha egy irreducibilis lánc egy bizonyos állapotba végtelen sokszor visszatér, akkor minden állapotba végtelen sokszor visszatér. (Ha ugyanis y egy másik állapot, akkor pozitív valószínűséggel elérhető y x -ből. Ha az x állapotban végtelen sokszor vagyunk, akkor végtelen sokszor lesz 0-tól különböző esélyünk arra, hogy az y állapotba jussunk. Ha egy eseménynek pozitív valószínűsége van, és végtelen sok kísérletet hajtunk végre, akkor ez az esemény 1 valószínűséggel végtelen sokszor be fog következni.)

Ha a lánc nem rekurrens, akkor minden állapot csak véges sokszor fordul elő. Ebben az esetben azt mondjuk, hogy a lánc *tranziens*. Gyakran nem könnyű feladat azt eldönteni, hogy egy Markov-lánc rekurrens vagy tranziens.

Tekintsük az $x \in S$ állapotot, és legyen $X_0 = x$. Jelölje R az x -be tett látogatások teljes számát, beleértve a kezdeti állapotot is. Ekkor

$$R = \sum_{n=0}^{\infty} I(X_n = x).$$

Itt I az indikátorfüggvényt jelöli, amely 1, ha az esemény bekövetkezik, egyébként pedig 0. Ha a lánc rekurrens, akkor R végtelen, ha pedig tranziens, akkor R értéke 1 valószínűséggel véges. Határozzuk meg R várható értékét:

$$E(R) = E \left[\sum_{n=0}^{\infty} I(X_n = x) \right] = \sum_{n=0}^{\infty} P(X_n = x) = \sum_{n=0}^{\infty} p_n(x, x).$$

Határozzuk meg most $E(R)$ -t más módon. Legyen T az x állapotba való első visszatérésig eltelt idő:

$$T = \min \{n > 0 : X_n = x\}.$$

Ha a lánc sohasem tér vissza x -be, akkor azt mondjuk, hogy $T = \infty$. Tegyük fel, hogy $P(T < \infty) = 1$. Ekkor 1 valószínűséggel a lánc mindig visszatér, és 1 valószínűséggel végtelen sokszor tér vissza, tehát a lánc rekurrens.

Tegyük fel most azt, hogy $P(T < \infty) = q < 1$, és határozzuk meg R eloszlását. Először is, $R = 1$ akkor és csak akkor, ha a lánc sohasem tér vissza, vagyis $P(R = 1) = 1 - q$. Legyen $m > 1$, ebben az esetben $R = m$ akkor és csak akkor, ha a lánc $m - 1$ -szer visszatér, de m -szer már nem, vagyis $P(R = m) = q^{m-1}(1 - q)$. Határozzuk meg R várható értékét, kihasználva, hogy $q < 1$:

$$\begin{aligned} E(R) &= \sum_{m=1}^{\infty} mP(R = m) = \sum_{m=1}^{\infty} mq^{m-1}(1 - q) = (1 - q) \sum_{m=1}^{\infty} mq^{m-1} \\ &= (1 - q) \frac{1}{(1 - q)^2} = \frac{1}{1 - q} < \infty. \end{aligned}$$

A várható értékre kapott két eredmény alapján kijelenthetjük a következőt:

Egy irreducibilis Markov-lánc akkor és csak akkor tranziens, ha egy állapotba való visszatérések számának várható értéke véges, vagyis akkor és csak akkor, ha

$$\sum_{n=0}^{\infty} p_n(x, x) < \infty.$$

Példa

Tekintsük a szimmetrikus véletlen bolyongást $\mathbb{Z}^d$ -ben. Mivel bármelyik állapot elérhető bármelyikből, ezért irreducibilis Markov-lánchról van szó, tehát a visszatérőséget elég egy állapotra vizsgálni. Az egyszerűség kedvéért legyen ez a 0 állapot. A visszatérőség vizsgálatához most szükségünk lesz az analízisből ismert STIRLING-formulára, mely szerint

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

$A \sim$ jelentése aszimptotikusan egyenlő, vagyis a két kifejezés hányadosa 1-hez tart, ha $n \rightarrow \infty$. Ez persze nem jelenti azt, hogy a különbségük 0-hoz tartana, sőt. Például $n = 100$ esetén a hányados értéke kb. 1,0008, a különbségük pedig kb. $7,7 \cdot 10^{154}$, $n = 500$ esetén pedig a hányados kb. 1,0002, a különbségük pedig kb. $1,3 \cdot 10^{865}$.

1. Legyen először $d = 1$. Ekkor az átmenetvalószínűségek:

$$p(x, x + 1) = \frac{1}{2}, \quad p(x, x - 1) = \frac{1}{2}.$$

Vegyük észre, hogy a visszatéréshez mindkét (bal-jobb) irányba azonos számút kell lépni, vagyis visszatérés csak páros számú, $2n$ lépésben lehetséges. Annak valószínűsége, hogy $2n$ lépésben visszatérünk:

$$p_{2n}(0,0) = \binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n = \binom{2n}{n} \left(\frac{1}{2}\right)^{2n} = \binom{2n}{n} \frac{1}{4^n} = \frac{(2n)!}{n!n!} \frac{1}{4^n}.$$

Az itt szereplő faktoriálisokat a Stirling-formula segítségével becsüljük. Ezzel:

$$\frac{(2n)!}{n!n!} \frac{1}{4^n} \approx \frac{\sqrt{2\pi 2n} \left(\frac{2n}{e}\right)^{2n}}{[\sqrt{2\pi n} \left(\frac{n}{e}\right)^n]^2} \frac{1}{4^n} = \frac{2\sqrt{\pi n} 4^n n^{2n} e^{-2n}}{2\pi n n^{2n} e^{-2n}} \frac{1}{4^n} = \frac{1}{\sqrt{\pi n}}.$$

Azt kaptuk tehát, hogy

$$p_{2n}(0,0) \sim \frac{1}{\sqrt{\pi n}}.$$

Mivel $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = \infty$, ezért

$$\sum_{n=0}^{\infty} p_{2n}(0,0) = \infty,$$

ami annyit jelent, hogy a lánc rekurrens (visszatérő).

2. Legyen most $d = 2$. Vegyük észre, hogy a visszatéréshez mindkét koordináta szerint vissza kell térni. Mivel a koordináták függetlenek, ezért ez az eset a

$$\left[\binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \right]^2$$

kifejezés vizsgálatát igényli. Az előzőek alapján

$$\left[\binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \right]^2 \sim \frac{1}{\pi n}.$$

Mivel $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$, ezért

$$\sum_{n=0}^{\infty} p_{2n}(0,0) = \infty,$$

ami annyit jelent, hogy a lánc most is rekurrens.

3. Legyen $d = 3$. A visszatéréshez mind a három koordináta szerint vissza kell térni. Mivel a koordináták függetlenek, ezért ez az eset a

$$\left[\binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \right]^3$$

kifejezés vizsgálatára vezet. Az előzőek alapján

$$\left[\binom{2n}{n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^n \right]^3 \sim \frac{1}{(\pi n)^{\frac{3}{2}}}.$$

Most viszont $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}} < \infty$, ezért ¹

$$\sum_{n=0}^{\infty} p_{2n}(0, 0) < \infty,$$

ami annyit jelent, hogy a lánc most tranziens (nem visszatérő).

A kapott eredményeket PÓLYA GYÖRGY² nevezetes tételében foglalhatjuk össze:

A szimmetrikus véletlen bolyongás $\mathbb{Z}^d$ -ben $d = 1$ és $d = 2$ esetén visszatérő, $d \geq 3$ esetén pedig nem visszatérő.

3 Pozitív rekurrens és nullrekurrens láncok

Legyen X_n egy irreducibilis, aperiodikus Markov-lánc megszámlálhatóan végtelen S állapottérrel. Azt fogjuk vizsgálni, hogy mely esetekben létezik határeloszlás. A $\pi(x)$, $x \in S$ határvalószínűségek valószínűség eloszlást határoznak meg S -en, úgy hogy minden $x, y \in S$ esetén

$$\lim_{n \rightarrow \infty} p_n(y, x) = \pi(x).$$

Ha X_n tranziens, akkor

$$\lim_{n \rightarrow \infty} p_n(y, x) = 0,$$

minden $x, y \in S$ esetén, azaz nem létezik határeloszlás. Létezhet viszont olyan rekurrens lánc, melyre a fenti határérték szintén 0. Tekintsük például a szimmetrikus véletlen bolyongást az egész számokon (ez éppen periodikus lánc, de nem baj). Ez mint láttuk rekurrens, de $p_{2n}(0, 0) \rightarrow 0$, ha $n \rightarrow \infty$. Egy Markov-lánc *nullrekurrens*, ha rekurrens, de

$$\lim_{n \rightarrow \infty} p_n(y, x) = 0.$$

¹A $\sum_{n=1}^{\infty} \frac{1}{n^{1+c}}$ végtelen sor $c > 0$ esetén konvergens.

²Magyar matematikus, tanítványa volt pl. Neumann János is.

Egyébként a rekurrens lánc *pozitív rekurrens*. A pozitív rekurrens láncok nagyon hasonló viselkedést mutatnak, mint a véges állapotterű láncok. Ha X_n irreducibilis, aperiodikus, pozitív rekurrens Markov-lánc, akkor minden $x, y \in S$ esetén a

$$\lim_{n \rightarrow \infty} p_n(y, x) = \pi(x) > 0$$

határérték létezik, és független y -től. A $\pi(x)$ -ek invariáns valószínűség eloszlást határoznak meg az állapottéren, azaz

$$\sum_{y \in S} \pi(y) p(y, x) = \pi(x) .$$

Továbbá, ha pl. az x állapotba való visszatéréshez szükséges időt vizsgáljuk,

$$T = \min \{n > 0 | X_n = x\} ,$$

akkor pozitív rekurrens lánc esetén

$$E(T | X_0 = x) = \frac{1}{\pi(x)} .$$

Ha X_n nullrekurrens, akkor $T < \infty$ 1 valószínűséggel, de $E(T) = \infty$. Tranziens lánc esetén $T = \infty$ pozitív valószínűséggel.

Egy módszer annak eldöntésére, hogy egy lánc pozitív rekurrens vagy sem az, ha megpróbáljuk megtalálni az invariáns eloszlást. Bebizonyítható, hogy ha egy irreducibilis lánc pozitív rekurrens, akkor egyértelműen létezik invariáns eloszlás, míg ha a lánc nem pozitív rekurrens, akkor nem létezik invariáns eloszlás. Ez utóbbi esetben a lánc lehet nullrekurrens, vagy tranziens.

4 Bernoulli-folyamat

Legyen X_n olyan sztochasztikus folyamat, amelyben X_n értéke az előzményektől függetlenül vagy 0, vagy 1. Ez persze így csak véges állapotterű folyamat, hiszen $S = \{0, 1\}$. Ennek ellenére érdemes vele foglalkozni, mert a belőle származtatott folyamatok és eloszlások igen fontosak a diszkrét idejű tömegkiszolgálási modellekben, másrészt nagyon hasonló tulajdonságokkal bír, mint a folytonos idejű megfelelője, a későbbiekben tárgyalandó Poisson-folyamat.

Legyen X_n eloszlása a következő (indikátor eloszlás, karakterisztikus eloszlás, Bernoulli-eloszlás):

$$X_n : \begin{cases} 0 & 1 \\ 1-p & p \end{cases}$$

Várható értéke és szórása:

$$E(X_n) = p \quad \text{és} \quad D(X_n) = \sqrt{p(1-p)}$$

A Bernoulli-folyamat egymástól független, Bernoulli-eloszlású valószínűségi változók sorozat. Mivel a lehetséges értékek 0 és 1, ezért egy lehetséges realizáció (pontosabban az eleje) például a következő:

$$0000110011110000011110101011 \dots$$

A 0 és 1 értékek helyett tekinthetünk mást is: sikeres-sikertelen, jött igény a tömegkiszolgáló rendszerbe vagy nem jött stb.

Milyen további kérdések merülnek itt fel?

- Adott n hosszúságú szakaszon mi a valószínűsége annak, hogy pontosan k darab 1-es (sikert) látunk?

A siker (1) valószínűsége függetlenül az előzményektől mindig p , így n kísérlet esetén annak a valószínűsége, hogy pontosan k darab sikereset látunk:

$$P(n\text{-ből pontosan } k \text{ sikeres}) = \binom{n}{k} p^k (1-p)^{n-k}$$

azaz a sikeresek száma binomiális eloszlással írható le.

- Mi a valószínűsége annak, hogy pontosan k kísérlet kell a következő sikerhez?

Mivel a kísérletek kimenetelei egymástól függtelenek, ezért ezt egy geometriai eloszlású valószínűségi változóval jellemezhetjük ($k-1$ darab sikertelen után a következő sikeres):

$$P(k \text{ kísérlet kell a következő sikerhez}) = (1-p)^{k-1} p$$

- Mennyit kell várnunk a következő sikeres kísérletre? Azaz: mennyi az első (vagy a következő) sikeres kísérletig tett próbálkozások számának várható értéke?

Mivel geometriai eloszlásról van szó, ezért a várható érték:

$$E(\text{következő sikeresig tett kísérletek száma}) = \frac{1}{p}$$

- Tegyük fel, hogy már egy ideje várunk egy sikeres kísérletre, mi a valószínűsége annak, hogy még legalább k darab kísérletre lesz szükségünk?

Jelölje Z az első sikeres kísérletig tett próbálkozások számát. Ekkor Z geometriai eloszlású valószínűségi változó. A kérdést pedig feltételes valószínűség formájában írhatjuk fel:

$$\begin{aligned} P(Z \geq k + r \mid Z > r) &= \frac{P(Z \geq k + r)}{P(Z > r)} = \frac{(1 - p)^{k+r-1}}{(1 - p)^r} \\ &= (1 - p)^{k-1} = P(Z \geq k) \end{aligned}$$

Az átalakítások során azt használtuk fel, hogy

$$\begin{aligned} P(Z > r) &= P(\text{az első } r \text{ sikertelen}) = (1 - p)^r \\ P(Z \geq k) &= P(\text{az első } k - 1 \text{ sikertelen}) = (1 - p)^{k-1} \end{aligned}$$

Azt kaptuk tehát, hogy

$$P(Z \geq k + r \mid Z > r) = P(Z \geq k)$$

Vagyis a sikertelen kísérletsorozat hosszának (r -nek) nincs szerepe a keresett valószínűségben. Másként megfogalmazva: az eloszlás emlékezet nélküli, örökifjú.

Melyik folytonos eloszlásnak vannak hasonló tulajdonságai?